UNIVERSITI TEKNOLOGI MARA

DESIGN AND DEVELOPMENT OF MOBILE CRIME RISK ASSESSMENT SYSTEMS (ICRAS) FOR INVESTIGATION OFFICERS USING ETL-BASED DATA PROCESSING AND PREDICTIVE ANALYSIS

MOHD RAZIF BIN MAIDIN

PhD

June 2026

UNIVERSITI TEKNOLOGI MARA

DESIGN AND DEVELOPMENT OF MOBILE CRIME RISK ASSESSMENT SYSTEMS (ICRAS) FOR INVESTIGATION OFFICERS USING ETL-BASED DATA PROCESSING AND PREDICTIVE ANALYSIS

MOHD RAZIF BIN MAIDIN

Thesis submitted in fulfilment
of the requirements for the degree of
Doctor of Philosophy
(Electrical Engineering)

Faculty of Electrical Engineering

June 2026

CONFIRMATION BY PANEL OF EXAMINERS

I certify that a Panel of Examiners has met on 23 February 2026 to conduct the final examination of Mohd Razif Bin Maidin on his Doctors of Philosophy thesis entitled "Design and Development of Mobile Crime Assessment Systems (ICRAS) for Investigation Officers Using ETL Based Data Processing and Predictive Analysis" in accordance with Universiti Teknologi MARA Act 1976 (Akta 173). The Panel of Examiners recommends that the student be awarded the relevant degree. The Panel of Examiners was as follows:

Norashikin M. Thamrin, PhD
Assosiate Professor
Faculty of Electrical Engineering
Universiti Teknologi MARA
(Chairman)

Ahmad Asari Sulaiman, PhD
Professor
Faculty of Electrical Engineering
Universiti Teknologi MARA
(Internal Examiner)

Ahmad Shah Hizam, PhD
Senior Lecturer
Faculty of Resilience
Rabdan Academy
(External Examiner)

PROFESSOR DR HJH ZURAEDA IBRAHIM
Dean
Institute of Postgraduates Studies
Universiti Teknologi MARA

Date: 12 June 2026

AUTHOR'S DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Teknologi MARA. It is original and is the results of my own work, unless otherwise indicated or acknowledged as referenced work. This thesis has not been submitted to any other academic institution or non-academic institution for any degree or qualification.

I, hereby, acknowledge that I have been supplied with the Academic Rules and Regulations for Post Graduate, Universiti Teknologi MARA, regulating the conduct of my study and research.

Name of Student	:	Mohd Razif Bin Maidin
Student ID. No.	:	2021405254
Programme	:	Doctor of Philosophy (Electrical Engineering) – CEEE950
Faculty	:	Electrical Engineering
Thesis Title	:	Design and Development of Mobile Crime Assessment Systems (ICRAS) for Investigation Officers Using ETL Based Data Processing and Predictive Analysis
Signature of Student	:	……..
Date	:	June 2026

ABSTRACT

Crime is a significant threat to public safety, social stability, and national development, particularly in rapidly urbanizing countries like Malaysia. Despite the Royal Malaysian Police (RMP) investing in crime reduction initiatives, the lack of a unified analytical framework hinders proactive crime prevention and resource optimization. Existing crime management strategies often rely on fragmented datasets, delayed reporting, and reactive decision-making, which limit early intervention efforts. This study aims to address these challenges by developing the Integrated Crime Risk Assessment System (ICRAS), a data-driven platform designed to enhance crime risk evaluation and management for law enforcement agencies. The primary objective is to consolidate multiple data streams, crime records, socio-economic indicators, spatial patterns, and environmental variables into a unified system for more effective crime risk assessment. The research employs an Extract, Transform, and Load (ETL) pipeline to clean, standardize, and merge the data, ensuring analytical accuracy and regional comparability. The system integrates two core analytical tools: the Risk Appetite Index and the Crime Risk Matrix, which quantify the likelihood and severity of crime incidents. To enable predictive capabilities, ICRAS utilizes the Random Forest algorithm, chosen for its robustness, interpretability, and ability to handle complex, nonlinear data. This model identifies key predictors of criminal activity and provides short-term forecasts to inform proactive policing. The dataset for this study includes 5,315 cases of violent and property crime in Malaysia from 2015 to 2021. Pilot testing conducted in Selangor showed that the Risk Appetite Index achieved a readability score of 4.52 and a Cronbach's α reliability coefficient of 0.86, while the Crime Risk Matrix scored 4.47 and 0.83, respectively. Both tools surpassed established thresholds for internal consistency and user comprehension. Furthermore, the implementation of ICRAS demonstrated improved investigative accuracy, reduced redundancy, and shortened decision-making cycles. Officers reported that the system's color-coded risk visualizations and structured descriptors enhanced interpretation, while its predictive insights supported faster, more confident judgments. ICRAS presents a methodologically sound and operationally viable solution for modern policing in Malaysia. By integrating ETL-based data processing, Random Forest prediction, and interactive visualization, this framework provides law enforcement agencies with a scalable, evidence-based tool for crime risk assessment. The findings not only offer a national solution but also present a transferable model for other jurisdictions seeking to institutionalize data-driven decision-making and enhance the analytical capabilities of their criminal justice systems.

ACKNOWLEDGEMENT

In Allah's Name, the most Merciful, all glory is to God, the Lord of the Worlds, and prayers and grace be to His servant and messenger Mohammad.

There are many individuals I would like to thank for their support and contributions in making the completion of this Thesis and PhD's Program a reality.

I am grateful to some people who have worked hard with me from the start to the completion of this study, particularly my supervisor, Associate Prof. Ts. Dr. Mohammad Fahmi bin Hussin @ Mohammad, my co-supervisor, Ts. Dr. Khairilmizal Samsudin, Ir. Ts. Dr. Nor Ayu Zalina Zakaria, and Ts. Dr. Amalina Binti Enche Ab. Rahim, who has always been generous throughout the research phases, and I wish to extend my gratitude to all of them for their constant encouragement to ensure that this study is completed

To my mother, to whom I owe a lot more than just thanks for the tolerance, endless encouragement, and support. I would also like to take this opportunity to express my warm regards to all my beloved family for their generous support throughout my life, and especially through my PhD process. I have the opportunity to finish this study because of their unconditional love and prayers.

To my loving wife, Norsuriyani Binti Hamzah, whose patience, support, and understanding throughout the past years have always encouraged me to take the road less traveled when I wanted to follow the pack, thank you. You are my best friend. To my loving daughter Medina Rawdha, my sons, Maqil Rayyan and Matin Reezq, who I hold dear to my heart.

TABLE OF CONTENTS

		Page
CONFIRMATION BY PANEL OF EXAMINERS		**ii**
AUTHOR'S DECLARATION		**iii**
ABSTRACT		**iv**
ACKNOWLEDGEMENT		**v**
TABLE OF CONTENTS		**vi**
LIST OF TABLES		**xi**
LIST OF FIGURES		**xiii**
LIST OF ABBREVIATIONS		**xvi**
CHAPTER 1 INTRODUCTION		**1**
1.1	Background	1
	1.1.1 Crime in Southeast Asia	3
	1.1.2 Crime in Malaysia	4
	1.1.3 Impact of Crime in Malaysia	6
	1.1.4 Crime Management in Malaysia	7
	1.1.5 Issues Facing Police in Malaysia	9
1.2	Problem Statement	13
1.3	Research Objective	14
1.4	Research Questions	16
1.5	Justification and Significance	16
1.6	Scope and Limitations	18
1.7	Thesis Outline	20
CHAPTER 2 LITERATURE REVIEW		**22**
2.1	Role of Technology	22
2.2	Crime Risk Assessment in Law Enforcement	24
2.3	Global Crime Risk Assessment Systems	27

2.4 Integrated Crime Risk Assessment Systems (ICRAS – Royal Malaysian Police) 31
2.5 Introduction to Crime Risk Assessment Models 32
2.5.1 Critical Evaluation of Existing Crime Risk Assessment Models 39
2.5.2 Analysis of Strengths and Weaknesses 41
2.5.3 Case Studies and Practical Applications 44
2.6 Predictive Modelling 49
2.6.1 Predictive Modelling in Law Enforcement 53
2.6.2 Predictive Modelling for Proactive Policing 55
2.6.3 Review of Studies on the Effectiveness of Predictive Modelling 58
2.7 Analysis of Case Studies and Real World Examples 60
2.8 Evaluation of Predictive Modelling Techniques 63
2.8.1 Selection of Predictive Model for ICRAS 69
2.9 Database Design in Crime Risk Assessment 70
2.9.1 The Importance of Data Processing in Crime Risk Assessment and Its Framework 72
2.9.2 Principles of Database Design in Crime Analysis 74
2.9.3 Database Design Principles 75
2.9.4 Data Sensitivity and Secrecy in Predictive Policing 77
2.9.5 Human Factors in Crime Prevention 78
2.10 Similar Implementation of Technology 79
2.11 Critical Gap Analysis 80

CHAPTER 3 RESEARCH METHODOLOGY 83
3.1 Research Design 83
3.1.1 Mixed Methods Research Design 85
3.1.2 Research Design of ICRAS 87
3.1.3 Research Phase 90
3.1.4 Integration of Data 94
3.2 Population and Sampling 98
3.2.1 Definition of Population 98
3.2.2 Data Collection Procedures 102
3.2.3 Sampling Procedure 103

3.2.4 Data Processing Using the Extract, Transform, and Load (ETL) Framework 105
3.2.5 Data Analysis and Interpretation 107
3.2.6 Methodological Considerations and Challenges 108
3.3 Quantitative Data Collection 109
3.3.1 Sources of Quantitative Data 109
3.3.2 Crime Data Extraction, Cleaning, and Processing 110
3.3.3 Quantitative Data Collection 112
3.3.4 Data Management Systems 114
3.3.5 Key Components of the Consent Process 116
3.3.6 Privacy and Anonymity 117
3.3.7 Compliance with Data Protection Laws 118
3.4 Data Analysis 119
3.4.1 Overview of Thematic Analysis 121
3.4.2 Integration of Statistical and Thematic Analysis 123
3.4.3 Predictive Modeling Using Random Forest Algorithm 124
3.4.4 System Challenges 127
3.5 Entity Relationship Diagram (ERD) and Data Model 127
3.5.1 Entities and Attributes 129
3.5.2 Entity Relationship Structure and Analytical Framework 130
3.6 Programming Languages and Frameworks 131
3.6.1 Database: MariaDB 10 134
3.6.2 Swagger API for Integration and Documentation 135
3.7 Data Flow and Process in ICRAS 135
3.7.1 Security and Data Privacy Considerations 137
3.8 Pilot Study 139
3.8.1 Overview of the Pilot Study 139

CHAPTER 4 RESULT AND DISCUSSION 144
4.1 Summary of Results 144
4.2 Data Collection Method 145
4.2.1 Questionnaire 145
4.2.2 Survey 146
4.2.3 Interview Data 148

4.2.4 Results of the ETL Process 150
4.3 Identifying the Need for ICRAS 153
4.4 Development of the Risk Appetite Index and Crime Risk Matrix 155
4.4.1 Determining Risk Appetite 155
4.4.2 Crime Risk Matrix 157
4.4.3 Data Sources of Crime Risk Matrix 160
4.4.4 Justification of the Matrix 161
4.4.5 Readability and Reliability of the Risk Appetite Index and Crime Risk Matrix 162
4.5 Quantitative Findings: Risk Scoring 164
4.6 Entity Relationships Diagram (ERD) 168
4.7 Development of the ICRAS System 186
4.7.1 System Design and Architecture 186
4.7.2 System Design Workflow 187
4.7.3 System Development 189
4.7.4 Authentication and Risk Assessment Flow User Process in ICRAS 193
4.8 Implementation of the ICRAS Framework 195
4.8.1 Machine Learning Performance in Crime Prediction 195
4.8.2 Pilot Testing of the ICRAS Framework 198
4.9 Real Time Crime Data Access and Risk Analysis in ICRAS 200
4.9.1 System Usability and Performance Evaluation 200
4.9.2 Impact on Crime Prevention Strategies 201
4.9.3 Integration into Law Enforcement Practices 203
4.10 Concluding Remark 205

CHAPTER 5 CONCLUSION 207
5.1 Summary of Key Findings 207
5.2 Contributions to Knowledge and Practice 208
5.3 Recommendations for Future Research 209

REFERENCES **213**

APPENDICES **233**

AUTHOR'S PROFILE **251**

LIST OF TABLES

Tables	Title	Page
Table 1.1	Crime Rates in Southeast Asia	4
Table 1.2	Summary of RMP Officers Killed/Injured in the Line of Duty Recorded in Malaysia	11
Table 2.1	Summary and Critical Evaluation of Selected Crime Risk Assessment Systems: Country of Use, Target Population, Model Type, Purpose, and Key Benefits and Limitations	30
Table 2.2	Strengths and Weaknesses of Crime Risk Assessment Models in Law Enforcement	43
Table 2.3	Comparison of Predictive Policing Models and Their Impact on Crime Reduction	63
Table 2.4	Comparative Analysis of Selected Machine Learning Algorithms in Crime Prediction	69
Table 2.5	Examples of Technological Innovations in Modern Law Enforcement	80
Table 2.6	Critical Gap Analysis of Existing Crime Risk Assessment Systems vs. Proposed ICRAS	81
Table 3.1	Phases of Crime Data Research: Quantitative and Qualitative Approaches	93
Table 3.2	Temporal Stratification of Crime Incidents by Year (2015–2021)	101
Table 3.3	Comparison of Random Forest and Deep Learning for Crime Risk Prediction	126
Table 3.4	Challenges and Solutions in the Development of ICRAS	127
Table 3.5	Entities and Relationships in ICRAS	129
Table 3.6	Relational Data Management in ICRAS using MariaDB	135
Table 3.7	Roles and Responsibilities of Participants in the ICRAS Pilot Study Evaluation	140
Table 3.8	Summary of Research Methodology for Integrated Crime Risk Assessment System (ICRAS)	141

Table 4.1	Preferences for Mobile Platform and Key Features in the Crime Risk Assessment System	145
Table 4.2	Summary of Survey Findings on Crime Risk Assessment Needs and ICRAS Expectations	147
Table 4.3	Summary of Interview Findings from Junior and Senior CID Investigation Officers on System Needs and Challenges	149
Table 4.4	ETL Process Comparison Table	153
Table 4.5	Summary of Validation, Readability, and Reliability Results for ICRAS Tools	163
Table 4.6	Summary of Statistical Results and Analysis for Crime Risk Matrix Validation	166
Table 4.7	Crime Incident Table	170
Table 4.8	Location Table	172
Table 4.9	Victims Table	173
Table 4.10	Crime Victims Table	175
Table 4.11	Suspect Table	176
Table 4.12	Crime Suspect Table	177
Table 4.13	Witness Table	179
Table 4.14	Crime Witness Table	180
Table 4.15	Evidence Table	181
Table 4.16	Officer Table	183
Table 4.17	Department Table	184
Table 4.18	ICRAS System Development Phases and Tools	188
Table 4.19	Machine Learning Results – Random Forest Framework	196
Table 4.20	Confusion Matrix Table	197
Table 4.21	Summary of Pilot Test Results for ICRAS	199

LIST OF FIGURES

Figures	Title	Page
Figure 1.1	Property & Violent Crime Statistic by Region [6]	1
Figure 1.2	Police Officers on Patrol, Representing Malaysia's Efforts in Community Policing and Adapting Law Enforcement Strategies to Combat Urban Crime	5
Figure 1.3	Police Act 1967 (Act 344), Regulation & Rules followed by Royal Malaysian Police (RMP) while Conducting the Investigation	10
Figure 1.4	Technology Integration in Crime Management and Law Enforcement	14
Figure 1.5	Advancing Crime Risk Assessment Through ICRAS Technology	18
Figure 2.1	Community Safety and Social Networking Apps – Citizen and Nextdoor	24
Figure 2.2	Illustration of Crime Risk Assessment Integrating Data Analytics, Community Input, and Strategic Policing to Identify Risks and Guide Preventive Actions	27
Figure 2.3	Visual Overview of Major Crime Risk Assessment Models Used Globally	29
Figure 2.4	RNR and VRAG Approach in Actuarial Mode	34
Figure 2.5	SAVRY and HCR-20	36
Figure 2.6	General Machine Learning Overview	39
Figure 2.7	Crime Reduction Trends	60
Figure 2.8	PredPol and SSL System used in the Real World	63
Figure 2.9	Workflow of Crime Prediction Using Machine Learning and Deep Learning Algorithms	66
Figure 3.1	Integration of Quantitative and Qualitative Insights for Enhanced Crime Risk Assessment and Decision Making	87
Figure 3.2	ICRAS Research Design Framework	88
Figure 3.3	Process Developments of ICRAS	91

Figure 3.4	Integrating Quantitative and Qualitative Data for Comprehensive Crime Risk Assessment	94
Figure 3.5	Mixed Methods Research Process	96
Figure 3.6	Comparison of Sampling Techniques	100
Figure 3.7	Overview of Crime Data Collection and Pre Processing Methodology	103
Figure 3.8	Systematic Sampling Framework for ICRAS Evaluation	105
Figure 3.9	ETL Based Data Processing Workflow for ICRAS	106
Figure 3.10	Comprehensive Frameworks for Data Analysis and Interpretation	107
Figure 3.11	Methodological Challenges and Key Considerations	108
Figure 3.12	Key Data Sources for Crime Analysis in ICRAS	110
Figure 3.13	Data Processing and Ethical Workflow for ICRAS Development	110
Figure 3.14	Inclusion Criteria for Stakeholder Participation in the ICRAS Study	113
Figure 3.15	Trauma Informed Techniques	117
Figure 3.16	Thematic Analysis Flowcharts	122
Figure 3.17	Entity Relationship Diagram (ERD)	129
Figure 3.18	Secure Backend Implementation Using PHP 8.2	132
Figure 3.19	ICRAS Process	136
Figure 3.20	Security and Data Privacy Elements	138
Figure 3.21	Overviews of Pilot Study Objectives for Evaluating ICRAS Performance and Usability	139
Figure 4.1	ETL Process Results Comparison	152
Figure 4.2	Processes for Determining Risk Appetite Index	157
Figure 4.3	ICRAS Risk Assessment Matrix (IRAM)	160
Figure 4.4	Simulated Dataset with Key Socio Economic and Policing Features	166
Figure 4.5	Train Test Split - Split the Data into Training and Testing Sets	167
Figure 4.6	Logistic Regression Model with Key Coefficients and ROC–AUC Performance	167
Figure 4.7	Decision Tree Model with Visualization and Feature Importance	167

Figure 4.8	Entity Relationships Diagram	168
Figure 4.9	Conceptual Layered Architecture of ICRAS	187
Figure 4.10	ICRAS Mobile Application Development Flowchart	189
Figure 4.11	ICRAS Main Dashboard Interface	192
Figure 4.12	Flowchart of ICRAS Mobile Apps	193
Figure 4.13	ICRAS Login Page	195
Figure 4.14	Risk Assessment and Scoring Interface	195
Figure 4.15	Performance Comparison Chart	201
Figure 4.16	Key Features of ICRAS for Investigation Officers	205

LIST OF ABBREVIATIONS

Abbreviations

AAB	Android App Bundle
AES	Advanced Encryption Standard
AES-256	Advanced Encryption Standard
AI	Artificial Intelligence
AJAX	Asynchronous JavaScript and XML
ANOVA	Analysis of Variance
APK	Android Package
ASEAN	Association of Southeast Asian Nations
ASEANAPOL	ASEAN Chiefs of Police
AUC	Area Under the Curve
CAS	Crime Anticipation System
CID	Criminal Investigation Department
CMA	Communications and Multimedia Act
CMI	Case Management Inventory
COMPAS	Correctional Offender Management Profiling for Alternative Sanctions
CPD	Chicago Police Department
CPTED	Crime Prevention Through Environmental Design
CRISP-DM	Cross-Industry Standard Process for Data Mining
CRM	Crime Risk Matrix

CSV	Comma-Separated Values
DNA	Deoxyribonucleic Acid
DOM	Document Object Model
DSS	Decision Support Systems
ERD	Entity Relationship Diagram
ETL	Extract, Transform, and Load
FDI	Foreign Direct Investment
FK	Foreign Key
GDPR	General Data Protection Regulation
GIS	Geographic Information System
HART	Harm Assessment Risk Tool
HCR - 20	Historical Clinical Risk Management 20
HIPAA	Health Insurance Portability and Accountability Act
HTML	HyperText Markup Language
ICRAS	Integrated Crime Risk Assessment System
INTERPOL	International Criminal Police Organization
IO	Investigation officers
IoT	Internet of Thing
IPCC	Independent Police Conduct Commission
IR 4.0	Fourth Industrial Revolution
IRAM	ICRAS Risk Assessment Matrix
IT	Information Technology
JIT	Just in Time

KDD	Knowledge Discovery in Databases
KDE	Kernel Density Estimation
LAPD	Los Angeles Police Department
LIME	Local Interpretable Model-Agnostic Explanations
LS	Level of Service
LSTM	Long Short-Term Memory
MCPF	Malaysia Crime Prevention Foundation
MFA	Multi-Factor Authentication
MLP	Multi-Layer Perceptron
NLP	Natural Language Processing
OASys	Offender Assessment System
OLAP	Online Analytical Processing
PDPA	Personal Data Protection Act
PHP	Hypertext Preprocessor
PK	Primary Key
PredPol	Predictive Policing
RAI	Risk Appetite Index
RBAC	Role-Based Access Control
RDMS	Research Data Management System
RMP	Royal Malaysian Police
RNN	Recurrent Neural Networks
RNR	risk, need, and responsivity
ROC	Receiver Operating Characteristic Curve

RTCC	Real Time Crime Centers
RTM	Risk Terrain Modelling
SAVRY	Violence Risk in Youth
SDLC	Software Development Life Cycle
SDSC	Chicago Police Department's Strategic Decision Support Centers
SHAP	SHapley Additive exPlanations
SPD	Seattle Police Department
SPSS	Statistical Package for the Social Sciences
SQL	Structured Query Language
SSL	Strategic Subjects List
SVM	Support Vector Machines
TLS	Transport Layer Security
UAT	User Acceptance Testing
UI	User Interface
UN	United Nation
UNODC	United Nations Office on Drugs and Crime
UX	User Experience
VARCHAR	Variable Character
VRAG	Violence Risk Appraisal Guide
WEF	World Economic Forum
XAI	Explainable Artificial Intelligence

CHAPTER 1
INTRODUCTION

1.1 Background

Crime is a complex and widespread social problem that affects communities worldwide, influencing public safety, economic stability, and social development [1]. The nature and occurrence of crime are shaped by multiple factors, including socio-economic disparities, urbanization, technological developments, and evolving criminal activities [2]. Crime encompasses a wide range of illegal activities, from violent offences such as homicide, assault, and robbery to property crimes including burglary, theft, and vandalism [3]. Although crime exists globally, its patterns and intensity vary across regions due to differences in socio-economic conditions and institutional effectiveness [4].

Global statistics indicate considerable variation in crime rates across regions. Latin America and North America report relatively high levels of property and violent crimes, while Europe and Oceania generally record lower crime rates. Asia reports comparatively lower crime rates overall, although rapid urbanization and economic development continue to introduce new crime-related challenges [1,5]. These patterns suggest that crime is often associated with socio-economic inequalities and structural conditions that influence opportunities for criminal activity [2,4].

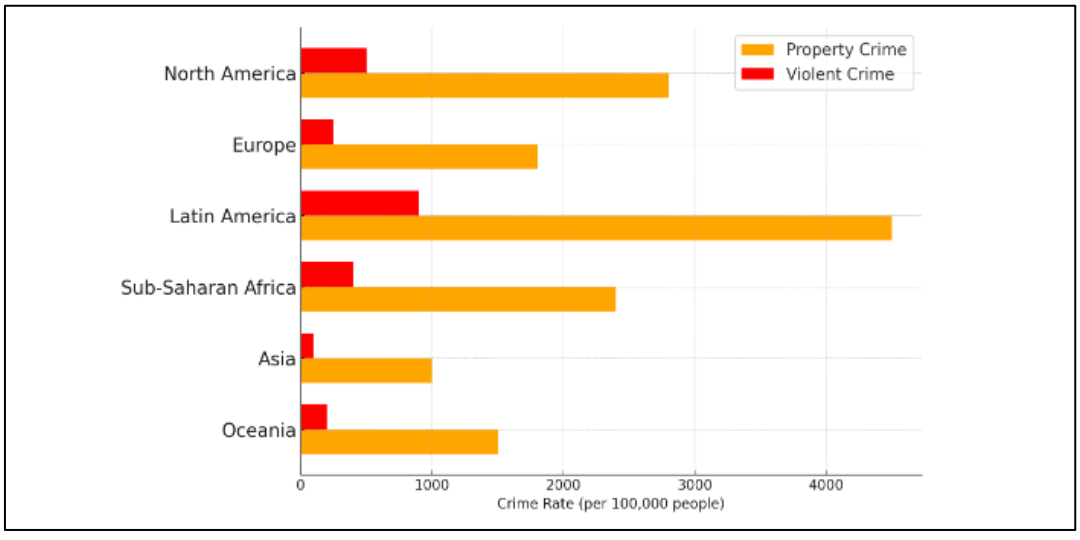


Figure 1.1 Property & Violent Crime Statistic by Region [6]

Figure 1.1 illustrates the distribution of property and violent crime rates across major world regions [6]. Property crimes such as burglary, theft, and motor vehicle theft are particularly prevalent in urban environments where population density and economic inequality create favourable conditions for criminal activity. According to the International Criminal Police Organization (Interpol), urban areas tend to experience higher levels of property crime due to the concentration of potential targets and increased anonymity for offenders [7]. The economic impact of property crimes is substantial, with financial losses amounting to billions of dollars annually [8]. The impact of crime extends beyond individual victims and affects broader social and economic development [9]. High crime rates can discourage investment, reduce tourism, and increase public expenditure on security and law enforcement. As a result, effective crime prevention strategies are essential for maintaining stable and secure communities. International organizations such as the World Economic Forum emphasize the importance of integrated approaches that combine law enforcement practices with technological innovation and data-driven decision-making [10]. Recent developments in crime prevention highlight the growing importance of analytical tools and information systems that enable law enforcement agencies to identify crime patterns and respond more effectively to emerging risks [11].

In this context, the use of data-driven approaches and predictive analytics has become increasingly important in supporting systematic crime risk assessment. Advanced analytical methods enable law enforcement agencies to transform large volumes of crime-related data into meaningful information that can support decision-making and resource allocation. The development of structured analytical frameworks and integrated information systems therefore represents an important step toward improving the effectiveness of modern crime prevention strategies. This research addresses these needs through the development of the Integrated Crime Risk Assessment System (ICRAS), a data-driven framework designed to support systematic crime risk assessment within the Royal Malaysian Police. By integrating crime data, analytical models, and structured risk evaluation into a unified platform, the proposed system aims to improve the accuracy, efficiency, and consistency of crime risk assessment. The study demonstrates how the integration of predictive modelling and structured risk evaluation can support more proactive and evidence-based crime prevention strategies.

1.1.1 Crime in Southeast Asia

Southeast Asia, a region characterized by rapid economic growth, urbanization, and increasing regional integration, presents a complex landscape of crime dynamics. Among the most pressing issues are property crimes, such as theft, burglary, and robbery, and violent crimes, including assault, homicide, and domestic violence. These crimes are influenced by a combination of socio-economic factors, cultural diversity, and cross-border interactions [12]. Urban areas experiencing rapid population growth, such as Bangkok, Jakarta, and Manila, is particularly vulnerable to property crimes, where opportunities for theft and burglary proliferate due to high population density, socio economic disparities, and insufficient deterrence mechanisms. Similarly, violent crimes in these urban centres are often linked to poverty, substance abuse, and interpersonal conflicts, disproportionately impacting marginalized communities. Studies indicate that poverty and income inequality, exacerbated by rapid urbanization, create environments where crime thrives, underscoring the need for comprehensive socio-economic interventions to address these root causes [13].

The geographic and economic diversity of Southeast Asia further complicates crime prevention efforts. The region's extensive coastlines and open borders facilitate transnational crimes that often intersect with local property and violent offenses. For instance, human trafficking networks frequently exploit individuals for forced labour or sexual exploitation, fuelling incidents of violence and theft. Regional conflicts and political instability in certain areas contribute to the proliferation of organized crime, which capitalizes on weak governance and law enforcement [14]. Despite these challenges, Southeast Asia has made significant strides toward combating crime through regional collaboration. The ASEAN Declaration on Transnational Crime, for example, emphasizes the importance of shared information, capacity building, and coordinated operations among member states. However, disparities in legal frameworks, resource availability, and enforcement capabilities across countries hinder the full realization of these collaborative efforts. Moreover, the fight against property and violent crimes in Southeast Asia demands a holistic and collaborative approach that integrates socio-economic reforms, community-based initiatives, and technological innovations. While regional efforts like those spearheaded by ASEAN provide a strong foundation for addressing crime, a more concerted effort is required to bridge gaps in capacity and coordination among member states. By addressing the structural

inequalities and economic disparities that fuel crime, Southeast Asia can move toward building safer, more resilient communities. This comprehensive approach underscores the need to view crime not merely as a legal issue but as a societal challenge requiring sustained, inclusive, and multidimensional solutions to ensure long term stability and security across the region.

Table 1.1
Crime Rates in Southeast Asia

Crime Type	Country	No of cases/rate (Year)	Source
Property Crime			
Burglary and Theft	Indonesia	Over 200,000 cases (2019)	ASEAN Stats,2019 [12]
	Philippines	Approximately 100,000 cases (2019)	ASEAN Stats,2019 [12]
Motor Vehicle Theft	Malaysia	Over 12,000 cases annually (recent years)	RMP,2020 [13]
Violent Crime			
Homicide	Philippines	6.5 per 100,000 population (2019)	UNODC,2020 [1]
	Thailand	Varies	
Assault	Thailand	Over 20,000 cases (2019)	ASEAN Stats,2019 [12]
Robbery	Thailand	High cases in urban centres like Bangkok (2019)	ASEAN Stats,2019 [12]
	Philippines	High cases in urban centres like Manila (2019)	ASEAN Stats,2019 [12]
Domestic Violence	Malaysia	Addressing with stricter laws and support mechanism	UN Women,2020 [13]

1.1.2 Crime in Malaysia

Malaysia, as a rapidly developing nation in Southeast Asia, encounters its own set of crime-related challenges. The country's economic development and urbanization have led to an increase in both property and violent crimes. The Royal Malaysian Police (RMP) reported a total of 83,475 crime cases in 2019, reflecting a slight decrease from the previous year, yet highlighting the persistent nature of crime [15]. The Malaysian government has implemented various initiatives to curb crime, including community policing and the use of technology in law enforcement. However, the complexity and evolving nature of crime require continuous adaptation and innovation in crime

prevention strategies. While the overall crime rate may have decreased slightly, Malaysia still faces significant challenges in combating various forms of criminal activities. Property crimes, such as theft and burglary, remain a concern, particularly in urban areas. Additionally, violent crimes, including robberies and assaults, continue to pose threats to public safety. To address these challenges, the Malaysian government has implemented various initiatives to curb crime. One such initiative is community policing, which involves law enforcement officers working closely with local communities to prevent and respond to crime. By fostering closer ties between the police and the community, this approach aims to enhance trust and cooperation, ultimately leading to more effective crime prevention and detection. Furthermore, the Malaysian government has embraced the use of technology in law enforcement as part of its crime prevention strategy. This includes the deployment of surveillance cameras in public areas, the use of data analytics to identify crime hotspots, and the development of mobile applications to facilitate reporting of crimes and emergencies. However, the complexity and evolving nature of crime require continuous adaptation and innovation in crime prevention strategies. As criminals become increasingly sophisticated in their methods, law enforcement agencies must stay ahead by leveraging technology, enhancing collaboration with communities, and investing in training and capacity building for their personnel. In conclusion, while Malaysia has made progress in addressing its crime-related challenges, there is still much work to be done. By adopting a multi-faceted approach that combines traditional policing methods with innovative strategies and technologies, Malaysia can further improve public safety and security for its citizens.



Figure 1.2 Police Officers on Patrol, Representing Malaysia's Efforts in Community Policing and Adapting Law Enforcement Strategies to Combat Urban Crime

1.1.3 Impact of Crime in Malaysia

Crime in Malaysia doesn't just affect the individuals involved it has a ripple effect that impacts everything from public perception of safety to the country's economic health. High crime rates can make investors think twice about putting their money into local businesses, and tourists may choose other destinations over Malaysia if safety is a concern. Even local businesses may struggle, often facing increased costs as they ramp up security measures. The Malaysia Crime Prevention Foundation (MCPF) emphasizes the strong link between crime rates and economic activity, especially as safety concerns can hinder the willingness of businesses and residents to fully invest and engage in local opportunities. (MCPF, 2020). One of the biggest impacts of crime is on foreign direct investment (FDI). Foreign investors typically look at stability, safety, and long-term prospects when choosing a location. When crime rates are high, especially in urban areas where crime can affect supply chains, transportation, and worker safety, investors may see it as a risky bet. This hesitation can lead to fewer new investments, with long-term effects on economic growth. [16] indicates that higher crime rates have a statistically significant impact on reducing FDI inflows in Malaysia, which could mean fewer jobs and slower economic development in high-crime areas. The government has made ongoing efforts to address crime as part of its broader economic development initiatives, but crime continues to pose a barrier to achieving its full economic potential. Tourism is another sector deeply affected by crime, as safety concerns can influence travellers' decisions on where to visit. While Malaysia is a country rich in cultural heritage and natural beauty, tourists are often influenced by global travel advisories, which may highlight crime risks. If tourists hear about incidents involving theft, scams, or violent crime in popular tourist areas, Malaysia's appeal as a destination may take a hit. [17] highlight how a single crime-related travel advisory can cause a noticeable dip in tourist numbers, underlining how perceptions of safety play a central role in sustaining tourism's contribution to the economy. Crime also raises the cost of living and doing business. Companies operating in high-crime areas often need to allocate more resources toward security measures, things like surveillance systems, private security staff, and insurance policies. These increased costs don't just cut into company profits; they're often passed down to consumers, making goods and services more expensive. A study by [18] focused on businesses in the Klang Valley and found that security expenses alone in higher-crime areas were about 15% more than in safer

areas. This increase affects everyone, from small business owners to everyday shoppers. Beyond the financial costs, the psychological effects of crime on communities are profound. Fear of crime can discourage people from going out, socializing, or even engaging with their neighbours. This erodes trust within communities, making public spaces feel less safe and reducing social cohesion. The Urban Well-being Report (2022) found that almost 70% of people living in high-density urban areas in Malaysia experience stress or anxiety due to crime, especially in lower-income neighbourhoods where resources to address safety concerns may be limited [19]. This lingering fear not only affects individuals but also weakens the community as a whole. To address the impact of crime on both the economy and public well-being, Malaysian policymakers have introduced initiatives like the Safe City Program and Crime Prevention Through Environmental Design (CPTED). These programs aim to deter crime by making urban spaces safer and more inviting, but researchers argue that we need a more comprehensive approach. For instance, addressing socio-economic inequalities through education, job opportunities, and support for under-resourced areas is essential for sustainable crime reduction. In short, reducing crime in Malaysia requires efforts on multiple fronts beyond policing and enforcement to address the broader social and economic drivers of crime [20]. In summary, crime in Malaysia has far-reaching effects on economic growth, public perception of safety, and overall quality of life. Reducing crime isn't just about making neighbourhoods feel safe; it's about creating an environment that encourages investment, boosts tourism, and fosters stronger communities. To achieve this, Malaysia needs to tackle both the immediate impacts of crime and the underlying issues that fuel it. This holistic approach, bringing together law enforcement, community initiatives, and socio-economic development, could pave the way for a safer and more prosperous Malaysia.

1.1.4 Crime Management in Malaysia

Crime management in Malaysia is fundamentally entrusted to the Royal Malaysian Police (RMP), with the Criminal Investigation Department (CID) serving as the primary department responsible for addressing criminal activities [21]. The CID operates as a specialized unit tasked with investigating, solving crimes, and ensuring offenders are brought to justice. This department handles a broad spectrum of criminal cases, ranging from property crimes and violent offenses to transnational criminal

activities. Its comprehensive approach encompasses the entirety of the crime management process, from the initial investigation to presenting cases in court, ensuring that justice is upheld while public safety is maintained [22]. The operational framework of the CID is meticulously designed to maximize efficiency in detecting and prosecuting criminal acts. The department's responsibilities span evidence gathering, forensic investigations, suspect interrogations, and witness interviews, alongside close collaboration with other law enforcement agencies. The integration of modern technologies, such as digital surveillance systems and advanced forensic tools, has significantly enhanced the CID's investigative capabilities. Research highlights the pivotal role of forensic science in criminal investigations, with techniques such as DNA analysis and digital forensics providing robust evidence that strengthens case outcomes. For instance, studies emphasize that the application of cutting-edge forensic tools has led to higher conviction rates and greater public confidence in law enforcement agencies globally. Within the CID, specialized divisions focus on particular categories of crime, such as homicide, sexual offenses, property crimes, and organized crime. This task-specific organization ensures expertise is channelled effectively, improving the department's capacity to address the unique challenges associated with different crime types. Scholars have noted the value of such specialization, as it allows for a more nuanced and expert-driven approach to tackling crimes, particularly in cases that demand in-depth technical knowledge or a tailored strategy.

The investigative process undertaken by the CID is both methodical and rigorous. Investigators begin with an in-depth examination of crime scenes, meticulously gathering physical evidence while piecing together eyewitness accounts. This evidence is supplemented by the analysis of digital data, such as surveillance footage and communication records. Forensic science plays a critical role throughout this process, ensuring that the evidence collected is both reliable and admissible in court. Studies on global law enforcement practices underscore the importance of these scientific methods in enhancing the accuracy and efficiency of crime solving. However, challenges persist. Resource constraints, an ever increasing caseload, and the dynamic nature of modern crimes such as cybercrime and financial fraud often test the department's resilience and adaptability. In addressing transnational crimes, the CID collaborates extensively with international law enforcement agencies and regional bodies such as INTERPOL and ASEANAPOL. These partnerships are vital for combating criminal networks that operate across borders, such as those involved in

human trafficking and drug smuggling. Joint operations and intelligence sharing initiatives enhance the CID's ability to disrupt these activities, ensuring a coordinated and comprehensive response. Research underscores the importance of such collaborative efforts, highlighting their effectiveness in dismantling complex criminal operations and ensuring global security. Central to the CID's mission is ensuring that offenders are held accountable through the judicial system. Once an investigation concludes, the CID works closely with prosecutorial authorities to present a compelling case against the accused. This partnership is critical to securing convictions and reinforcing public trust in the justice system. Studies have shown that effective collaboration between law enforcement and prosecutorial bodies is a key determinant of successful outcomes in criminal cases. By presenting evidence that is both substantial and credible, the CID plays a crucial role in upholding the integrity of Malaysia's judicial process.

1.1.5 Issues Facing Police in Malaysia

The Royal Malaysian Police play a critical role in maintaining law and order, preventing crime, and protecting citizens. The RMP's task force, including specialized units such as the Criminal Investigation Department (CID), focuses on various aspects of crime prevention and investigation. Investigation officers (IOs) in the CID face numerous challenges, including limited resources, increasing crime complexity, and the need for advanced training and technology to effectively combat modern crimes. At the same time, IOs must adhere to the Police Act 1967 (Act 344) and follow the regulations and rules set forth by the Royal Malaysian Police (RMP) while conducting investigations in Figure 1.3. These challenges include:

i. Limited Resources: The RMP often operates with limited manpower and financial resources, which can hinder the implementation of comprehensive crime prevention strategies.
ii. Technological Challenges: Rapid advancements in technology require continuous updates in law enforcement techniques and tools. Without adequate resources and training, the police may struggle to keep pace with technological developments used by criminals.

iii. Complexity of Crime: The nature of crime is constantly evolving, with criminals becoming more sophisticated in their methods. This demands advanced investigative skills and coordination among law enforcement agencies to effectively combat modern crimes.
iv. Public Trust: Building and maintaining public trust is crucial for effective policing. Incidents of corruption and misconduct within the police force can undermine public confidence and cooperation, making it harder to prevent and solve crimes [23].

To address these challenges and effectively combat crime, the RMP requires continuous support, both in terms of resources and public cooperation. Strengthening law enforcement capabilities, enhancing technological infrastructure, and fostering greater community engagement are essential for improving the safety and security of Malaysian society.

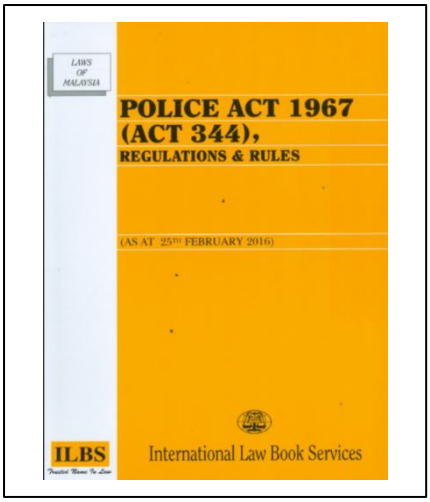


Figure 1.3 Police Act 1967 (Act 344), Regulation & Rules followed by Royal Malaysian Police (RMP) while Conducting the Investigation

In their role of ensuring national security and public safety, officers often face significant personal risks. As the primary law enforcement agency tasked with investigating crimes, apprehending offenders, and maintaining order, the RMP operates in a challenging environment that demands both expertise and resilience. Among the many sacrifices made by the men and women of the force, some have paid the ultimate price while fulfilling their duty, as shown in Table 1.2. Documenting and honouring these sacrifices not only highlights the dangers inherent in policing but also underscores

the commitment of law enforcement personnel to upholding justice. Officers in the RMP often face dangerous situations, ranging from armed confrontations with criminals to high-risk operations in volatile environments. These scenarios underscore the complexities of modern crime, including the proliferation of violent crimes such as armed robberies and transnational criminal activities. Global research in criminology has shown that law enforcement fatalities are often linked to direct encounters with offenders, vehicular accidents during pursuits, and exposure to hazardous materials during investigations. These risks are increased by the unpredictable nature of crime and the evolving tactics of criminal organizations. Furthermore, a comprehensive record of police injuries sustained during their service serves as a stark reminder of the inherent risks and sacrifices associated with law enforcement [24]. This recorded table not only highlights the physical dangers faced by officers in the line of duty but also underscores the unwavering dedication and commitment they demonstrate in ensuring public safety and security. By acknowledging the sacrifices made by law enforcement personnel, society can better appreciate the challenges they confront daily and work towards providing the necessary support and resources to safeguard their well-being.

Table 1.2
Summary of RMP Officers Killed/Injured in the Line of Duty Recorded in Malaysia

Year / Month	Location	Rank	Event type
2010, Mac	Sarawak	Cpl	Killed at the scene when Naza Citra MPV type boarding by three men failed to control before encountering and collided his police Land Rover.
2013, Mac	Sabah	Insp	Had been shoot by the terrorist group from Southern Philippines known as "Royal Security Forces of the Sultanate of Sulu and North Borneo".
2013, Mac	Sabah		Killed by the terrorist while investigate the house in Sri Jaya Village which believed the existence of a cache of firearms and Sulu terrorist hideout in the village off the coast of Semporna, Sabah.
2013, Sep	Malacca	Cpl	Killed in a struggle to detain a suspect in an operation, code-named "Ops Lejang".
2014, Mac	Selangor	Cpl	Stabbed to death by three Indonesians who were attempting to steal his weapon.
2014, Mei	Perak	L/Cpl	Motorcycle was collided from the rear by a Honda Civic driven by the robbers' accomplice.

Year / Month	Location	Rank	Event type
2017, Aug	Selangor	L/Cpl	Assaulted in close range by his assailants at the time as an officer had carried out an inspection at the station
2020, Mei	Selangor	Cpl	Killed by a drunk driver in Highway when conducting roadblock.
2020, Nov	Perlis	Cpl	Ambushed by ten or thirteen assailants at early morning, believed it is kratom smugglers. In the incident, officers died at the scene.
2021, Aug	Selangor	Sgt	Died as a result of an accident while chasing an illegal racing suspect in the Street Gangster Operation

Considering the escalating number of officers injured or killed in the line of duty in Malaysia, the imperative for effective crime management has become paramount. It is essential to bolster support for crime management initiatives in Malaysia, such as community policing, specialized task forces, intelligence gathering, and enhanced technology use, through comprehensive policies, frameworks, and mechanisms practiced by the Royal Malaysia Police (RMP). Additionally, global policies aimed at preventing such tragedies must be considered. One significant system in this regard is the Integrated Crime Risk Assessment System (ICRAS), which serves several objectives in enhancing law enforcement capabilities.

Despite the potential benefits of technology-driven approaches to crime risk assessment, many law enforcement agencies still lack a comprehensive and integrated system to support their investigative efforts [25]. This deficiency can hinder the efficiency and effectiveness of crime prevention and investigation efforts, leaving agencies without the necessary tools to effectively combat criminal activity. Implementing robust technology systems tailored to the needs of law enforcement can enhance data analysis, predictive modelling, and resource allocation, ultimately leading to more proactive and targeted policing strategies [26]. However, the adoption of such systems requires substantial investment in both technology infrastructure and training for personnel to ensure successful integration and utilization within law enforcement operations. Therefore, while technology offers promising solutions for improving crime risk assessment and investigation, addressing the existing gaps in system integration and capacity remains a crucial challenge for many law enforcement agencies. The fragmented nature of existing tools and data sources often hampers the ability of Investigation Officers (IOs) to make timely and well-informed decisions.

1.2 Problem Statement

Effective crime risk assessment is essential for law enforcement agencies as it supports the identification, analysis, and mitigation of threats to public safety and security [25]. However, the Royal Malaysian Police (RMP) currently faces significant challenges due to the absence of a dedicated and integrated system for systematic crime risk analysis. Crime risk assessment activities are largely conducted using fragmented and manual approaches, which limit the efficiency and effectiveness of crime prevention and resource allocation. A major technical limitation arises from the difficulty faced by Investigation Officers (IOs) in integrating data from multiple sources, including crime records, socio-economic indicators, and environmental factors. At present, IOs are required to manually collect, clean, and analyze data from different departments and information systems. This process is time-consuming and prone to human error, often requiring several hours or days, depending on the volume and complexity of the data. The lack of real-time data integration delays risk assessments and reduces the ability of law enforcement personnel to respond promptly to emerging threats. Furthermore, fragmented data storage across departments limits the ability to conduct comprehensive and consistent crime risk analysis.

In addition to technical limitations, Investigation Officers often have limited access to analytical tools and structured risk assessment methods. As a result, crime risk evaluations frequently depend on traditional approaches or individual experience, which may lead to inconsistencies in decision-making and difficulties in prioritizing high-risk cases. Operational constraints, including restricted access to up-to-date information and challenges in inter-departmental data sharing, further reduce the effectiveness of crime risk assessment processes. Consequently, the current approach remains largely reactive and provides limited support for proactive crime prevention and strategic resource planning.

The development of a computerized and integrated crime risk assessment system offers significant advantages over traditional manual approaches. Computerized systems provide continuous accessibility, consistent performance, and the ability to process large volumes of data efficiently. Unlike manual analysis, automated systems can integrate information from multiple sources, update data continuously, and provide standardized outputs for decision-making. Such systems also reduce dependence on individual expertise and minimize delays associated with manual data processing and

analysis. To address these limitations, there is a need for a technology-driven solution that integrates modern information technology tools, including data analytics, machine learning, and predictive modelling, to support systematic crime risk assessment [30]. Automating the data integration and analysis processes would reduce delays and improve the accuracy and reliability of crime risk evaluations. The availability of user-friendly interfaces and visualization tools would further enable Investigation Officers to interpret crime data more effectively and make timely decisions. Addressing the gaps in current crime risk assessment practices requires a comprehensive approach that combines technological innovation with improved analytical capabilities and data integration. By leveraging modern IT tools and structured analytical methods, law enforcement agencies can enhance their ability to anticipate crime risks, allocate resources effectively, and improve public safety outcomes [27]. This research, therefore aims to develop an integrated crime risk assessment framework that supports systematic and data-driven decision-making within the Royal Malaysian Police.



Figure 1.4 Technology Integration in Crime Management and Law Enforcement

1.3 Research Objective

This research introduces a novel, data-driven approach to crime risk assessment through the development of the Integrated Crime Risk Assessment System (ICRAS). Unlike traditional methods that rely on static crime data, ICRAS integrates real-time data from multiple sources, such as historical crime data, socio-economic factors, and geospatial data, enabling dynamic crime predictions. This system uses Random Forest algorithms to process complex relationships between various data points and predict emerging crime hotspots. This new approach is designed to shift from reactive policing

to proactive crime prevention, offering timely, actionable insights for law enforcement agencies. Specific goals related to ICRAS development and assessment of crime risk assessment practices include:

i) To introduce a risk appetite index and crime risk matrix for assessing the likelihood of criminal cases within the framework of the Royal Malaysian Police (RMP). This involves developing a comprehensive risk assessment framework that considers various factors such as crime patterns, geographical locations, demographic characteristics, and socio-economic conditions. The goal is to provide IOs with standardized tools and methodologies for evaluating the severity and probability of criminal incidents, allowing for more informed decision-making and resource allocation.
ii) To design and developed framework for the Integrated Crime Risk Assessment System (ICRAS) that integrates advanced data analytics techniques and machine learning algorithms. This framework will enable the aggregation and analysis of data from diverse sources, including law enforcement databases, social media feeds, and public records, to provide a holistic view of crime risks. The objective is to design and developed a user-friendly interface and dashboard that empowers IOs with intuitive tools for visualizing and interpreting crime data, facilitating more effective communication and collaboration among law enforcement agencies.
iii) Develop an ICRAS platform tailored for Investigation Officers, enabling real-time crime data access, risk analysis, and generation of actionable insights. This aims to streamline crime risk assessment, enhancing IOs' ability to manage risks and improve public safety in Malaysia.

By achieving these specific goals, the research aims to address the existing gaps and challenges in current crime risk assessment practices and contribute to the development of innovative solutions that enhance the capabilities of law enforcement agencies in preventing and combating criminal activities.

1.4 Research Questions

In the pursuit of enhancing crime risk assessment practices within law enforcement agencies, particularly among Investigation Officers (IOs) in Malaysia, this study aims to address critical challenges and limitations through a systematic investigation and the development of an Integrated Crime Risk Assessment System (ICRAS). The research questions guiding this endeavour include:

1) What are the key challenges and limitations in current crime risk assessment practices and existing systems used by law enforcement agencies, particularly IOs in Malaysia?
2) What innovative technological solutions, such as machine learning algorithms, can be employed to enhance the accuracy and reliability of crime risk assessments?
3) How do factors such as gaps in expertise and experience among IOs, and constraints in accessing pertinent information or data, increase efficiency in conducting thorough risk assessment?
4) What are the essential components of a comprehensive framework for the Integrated Crime Risk Assessment System (ICRAS), and how can these components be designed to integrate data from diverse sources effectively?
5) How can user-friendly interfaces and dashboards be developed to empower IOs with intuitive tools for visualizing and interpreting crime data, facilitating more effective communication and collaboration among law enforcement agencies?

These research questions serve as a roadmap for investigating the complexities of crime risk assessment and guiding the development and implementation of ICRAS, with the goal of improving public safety and security.

1.5 Justification and Significance

The justification and significance of this research work extend beyond the immediate challenges faced by law enforcement agencies in Malaysia, encompassing broader implications for the field of crime prevention and public safety. At its core, this

research seeks to address the pressing need for more effective crime risk assessment practices, which are essential for identifying and mitigating potential threats to society. By delving into the limitations of existing systems and proposing innovative solutions, this research aims to contribute to the advancement of crime prevention strategies and the enhancement of overall security measures. Human experts, who traditionally conduct crime risk assessments, often face limited availability as their working hours are typically scheduled, restricting continuous real-time monitoring. In contrast, computerized systems such as the proposed Integrated Crime Risk Assessment System (ICRAS) offer greater accessibility, operating around the clock without requiring extensive user training. Additionally, the physical presence of human experts is constrained by logistical challenges, as they cannot be at multiple locations simultaneously, leading to inefficiencies and delays caused by transportation. ICRAS overcomes this limitation by providing a ubiquitous platform accessible from various locations, enabling simultaneous and timely assessments regardless of geographical barriers.

Moreover, human experts are vulnerable to physical risks in the field, which may threaten their safety and disrupt assessment operations, whereas computerized systems remain resilient to such threats, ensuring durability and reliable performance. The knowledge of human experts is also bounded by their individual learning and experience, which may lag behind rapidly evolving crime patterns. By integrating data analytics and machine learning, ICRAS facilitates swift and continuous updates to its knowledge base, improving the accuracy and relevance of crime risk evaluations. Furthermore, the speed and efficiency of human experts can vary due to factors like fatigue and subjective judgment, while computerized systems deliver consistent and standardized performance, enhancing the quality and timeliness of decision-making. Finally, the high costs and time investment required to train new human experts pose significant challenges to law enforcement agencies. ICRAS offers a cost effective solution, enabling multiple users to access the system without repetitive training expenses, thereby optimizing resource allocation.

This research holds significant promise in addressing critical gaps and challenges in current crime risk assessment practices. Through the development and implementation of ICRAS, this research aims to not only improve the capabilities of law enforcement agencies in Malaysia but also contribute to broader efforts to enhance public safety and security. By advancing the state-of-the-art in crime prevention

methodologies, this research has the potential to create lasting positive impacts on communities and societies at large.

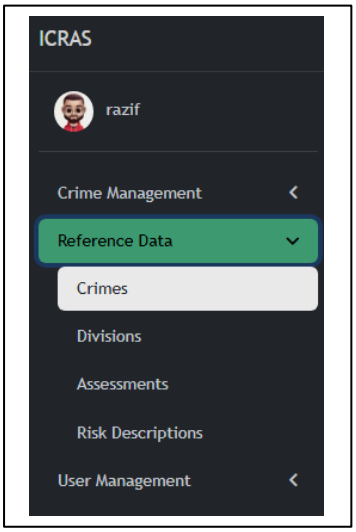


Figure 1.5 Advancing Crime Risk Assessment Through ICRAS Technology

1.6 Scope and Limitations

This research focuses on the technical design, development, and evaluation of the Integrated Crime Risk Assessment System (ICRAS), a data-driven platform intended to support crime risk assessment for Investigation Officers (IOs) within the Royal Malaysian Police (RMP). The scope of the study encompasses the development of a structured analytical framework that integrates crime-related datasets into a unified system capable of supporting systematic crime risk evaluation. The research includes the design of a backend data management structure and a user-accessible interface that enables Investigation Officers to retrieve and interpret crime risk information in a consistent and operationally meaningful manner. The primary emphasis of the study is on the technical components required to support crime risk assessment, including data integration, predictive modelling, and risk evaluation. The study incorporates structured datasets relevant to crime risk analysis, including crime incident records as well as selected socio-economic and operational indicators that influence crime patterns. Crime data consisting of 5,315 recorded cases of violent and property crimes between 2015 and 2021 were used to develop and evaluate the predictive and analytical components of the system. The geographical focus of the study is limited to the state of Selangor, which serves as the pilot region for the implementation and testing of the ICRAS

framework. This dataset provides a representative sample for system development and validation while allowing the research to be conducted within a controlled operational context.

A central component of the research involves the implementation of an Extract, Transform, and Load (ETL) framework to support the integration of data from multiple sources. The ETL process includes data extraction, cleaning, normalization, and consolidation into a unified database structure to ensure consistency and reliability for subsequent analysis. This data processing framework supports the analytical functions of ICRAS by enabling efficient data retrieval and standardized input for predictive modelling. The study also includes the development of two structured risk assessment instruments, namely the Risk Appetite Index and the Crime Risk Matrix, which provide standardized approaches for evaluating the likelihood and severity of crime risks. These instruments combine quantitative modelling outputs with contextual factors such as socio-economic conditions, geographic characteristics, and historical crime patterns to support consistent and interpretable risk assessments. To enhance the analytical capability of the system, the research integrates predictive modelling using the Random Forest algorithm to estimate crime risk levels based on selected predictor variables. The Random Forest approach was selected due to its robustness when applied to moderate-sized and heterogeneous datasets and its ability to produce interpretable results through feature importance analysis. The predictive model was trained and evaluated using the processed dataset, and its performance was assessed using established validation metrics, including accuracy, precision, recall, and Area Under the Receiver Operating Characteristic Curve (ROC-AUC). In addition to the analytical components, the research includes the development of user-oriented interfaces and visualization dashboards that enable Investigation Officers to interpret crime risk information in near real-time, thereby supporting faster and more informed decision-making.

The evaluation of the ICRAS framework is limited to a pilot implementation in Selangor, where the system was tested using historical crime data and user feedback from relevant stakeholders. While the pilot implementation provides meaningful insights into system performance and usability, the geographical scope limits the immediate generalization of the findings to other regions of Malaysia. Further validation using datasets from additional states would be required to establish broader national applicability. Several limitations are acknowledged in this research. The quality and completeness of crime data remain important constraints, as operational

datasets may contain missing or inconsistent records that can affect predictive accuracy despite the use of data cleaning and normalization procedures. In addition, the predictive modelling component is limited to the use of the Random Forest algorithm, and although this approach provides reliable and interpretable results, the investigation of more data-intensive modelling techniques such as deep learning is beyond the primary scope of the study. Despite these limitations, the research establishes a structured and scalable framework for crime risk assessment by integrating ETL-based data processing, predictive modelling, and standardized risk evaluation into a unified platform. The Selangor pilot implementation therefore serves as a proof of concept and provides a foundation for the future expansion of the ICRAS framework to other regions and operational environments.

1.7 Thesis Outline

The structure of the thesis is organized into six distinct chapters, each serving a specific purpose in the exploration and presentation of the research findings.

Chapter 1 serves as the introduction, providing an overview of the research topic, including its background and context. It also outlines the problem statement, research objectives, questions, and the justification for the study's significance. The scope and limitations of the research are delineated, along with a detailed explanation of the chosen research methodology. The chapter concludes with an outline of the entire thesis, summarizing the content of each subsequent chapter.

Chapter 2 comprises the literature review, where existing knowledge and research on crime risk assessment are synthesized and critically analysed. This chapter examines various crime risk assessment models, identifies challenges and limitations in current practices, explores the integration of technology in crime prevention, discusses ethical and legal implications, and highlights the role of human factors in crime prevention. It also includes a summary of the literature review and identifies gaps for further investigation.

Chapter 3 details the research methodology employed in the study. It describes the research design, including the rationale for choosing mixed-methods approach, and explains the population and sampling techniques. The chapter outlines the data collection methods, both quantitative and qualitative, along with the data analysis techniques used. It also provides insights into the development and testing of the

Integrated Crime Risk Assessment System (ICRAS) and addresses ethical considerations.

Chapter 4 presents the results and discussion of the study. It provides an in-depth analysis of the findings obtained from both quantitative and qualitative data collection methods. The chapter discusses the effectiveness of the Model of Incident Potential Index and the ICRAS Database, highlights key trends, evaluates the usability and accuracy of ICRAS, and compares findings with existing crime risk assessment models. Additionally, the chapter includes feedback from law enforcement officers and experts, as well as the implications of the findings on crime prevention strategies.

Chapter 5 serves as the conclusion of the thesis. It summarizes the key findings, discusses their implications, and reflects on the contributions of the study to the field of crime risk assessment. The chapter also addresses the limitations of the research and provides recommendations for future studies and potential improvements to the ICRAS framework. Finally, it highlights the overall significance of the research in enhancing law enforcement strategies and crime prevention efforts.

CHAPTER 2
LITERATURE REVIEW

2.1 Role of Technology

Technology has become the backbone of modern crime prevention strategies, fundamentally changing how communities and law enforcement agencies interact. Mobile applications, in particular, have emerged as powerful tools for fostering community engagement and enhancing crime reporting efforts. Apps like Citizen and Nextdoor exemplify this trend, enabling citizens to report suspicious activities and incidents in real time. Research shows that these platforms help bridge the gap between law enforcement and community members, encouraging collaboration and empowering individuals to take an active role in safeguarding their neighbourhoods [28]. These mobile applications create a two-way communication channel, allowing law enforcement to deliver timely updates, safety alerts, and crime prevention tips directly to residents. This immediacy not only raises public awareness but also builds a sense of unity and collective responsibility. Community members can share insights and experiences, enrich the overall understanding of local crime dynamics and strengthen the community's resilience against potential threats [29]. Digital platforms extend the reach and impact of crime prevention initiatives even further. Social media channels, such as Facebook and Twitter, provide law enforcement agencies with powerful tools for disseminating information quickly. By educating the public on crime trends, safety protocols, and the importance of community vigilance, these platforms enhance transparency and foster trust between law enforcement and the community [30]. Engaging with residents through social media can mobilize communities around specific issues, encouraging participation in neighbourhood watch programs or community safety events. Studies indicate that neighbourhoods with active law enforcement engagement on social media often experience lower crime rates, as this enhanced communication nurtures a sense of partnership and accountability [31]. The integration of technology in crime prevention has fundamentally transformed the landscape of law enforcement, providing agencies with sophisticated tools to enhance their operational efficiency and effectiveness. Among the most impactful innovations are mobile applications and digital platforms, which have emerged as vital resources

for both law enforcement and community engagement. Mobile applications, such as crime reporting tools and safety alert systems, empower citizens to report incidents in real time, fostering greater collaboration between the community and law enforcement agencies [32]. Research indicates that these technologies not only facilitate immediate communication but also allow for the aggregation and analysis of crime data, helping to identify patterns and hotspots that inform strategic deployment of resources [33]. Furthermore, digital platforms enable law enforcement to disseminate information swiftly, enhancing public awareness and encouraging community vigilance, which are critical elements in crime deterrence [34]. Moreover, the use of predictive analytics in conjunction with these technologies enables law enforcement agencies to anticipate potential criminal activities and allocate resources proactively. By harnessing the power of big data, agencies can develop comprehensive crime prevention strategies tailored to specific community needs [35]. The integration of technology into crime prevention is not merely an enhancement of traditional methods; it represents a paradigm shift that redefines how law enforcement interacts with communities and addresses crime. Technological advancements have profoundly transformed law enforcement capabilities, significantly enhancing crime prevention and investigative processes. One of the most significant developments in this realm is the adoption of predictive policing algorithms, which utilize machine learning and sophisticated data analytics to forecast criminal activities. By analysing historical crime data, demographic trends, and environmental factors, these algorithms provide law enforcement agencies with actionable insights that help them proactively address potential crime hotspots. Research shows that agencies employing predictive policing have seen reductions in crime rates, as these tools enable a more efficient allocation of resources and targeted interventions [36, 37]. In addition to predictive policing, advancements in forensic technology have revolutionized crime scene investigations. Techniques such as DNA analysis have become standard practice, allowing law enforcement to identify suspects with remarkable accuracy and speed. The National Institute of Justice reports that modern DNA profiling can match samples with an accuracy rate exceeding 99.9%, significantly increasing the likelihood of successful prosecutions [38]. Furthermore, technologies like aerial surveillance using drones and high resolution cameras enable law enforcement to monitor areas more effectively, gathering crucial evidence while minimizing the risk to personnel. Moreover, advancements in digital forensics have enhanced the ability of law enforcement agencies to recover and analyse electronic

evidence. This includes everything from extracting data from smartphones to analysing digital footprints left in cyberspace. Studies have shown that effective digital forensics can lead to more efficient investigations and a higher rate of case resolutions [39].

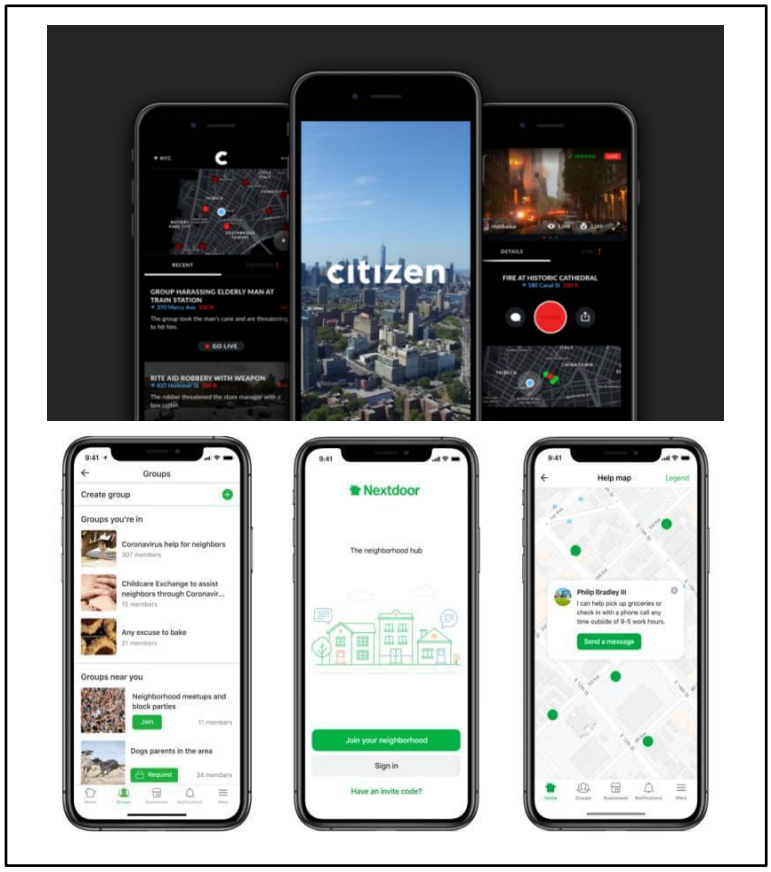


Figure 2.1 Community Safety and Social Networking Apps – Citizen and Nextdoor

2.2 Crime Risk Assessment in Law Enforcement

Crime risk assessment can be defined as the systematic process of evaluating and analysing various factors that contribute to criminal activities within specific regions or communities. These factors may include but are not limited to crime patterns, geographical locations, socio-economic conditions, demographic characteristics, and environmental factors [40]. By dissecting and assessing these elements, law enforcement agencies can gain valuable insights into crime, enabling them to identify high risk areas, vulnerable populations, and emerging crime trends. The significance of crime risk assessment in law enforcement lies in its ability to inform evidence based policing practices and strategic crime prevention initiatives. By optimizing data driven insights and analytical tools, law enforcement agencies can develop a comprehensive

understanding of crime dynamics, enabling them to prioritize interventions, allocate resources efficiently, and deploy personnel strategically [41]. Moreover, crime risk assessment facilitates collaboration and information-sharing among various stakeholders, including law enforcement agencies, government entities, community organizations, and academic institutions [42]. Assessing and managing crime risks is essential for ensuring public safety and security, as it enables law enforcement agencies to identify potential threats and vulnerabilities before they escalate into significant security concerns. By proactively addressing root causes and underlying risk factors contributing to criminal activities, agencies can implement preventive measures and enforcement strategies effectively. Moreover, crime risk assessment guides investigative efforts by identifying potential suspects, modus operandi, and patterns of criminal behaviour, enabling law enforcement agencies to focus their resources on cases with the highest likelihood of resolution and successful prosecution [43]. In practical terms, crime risk assessment informs a wide range of law enforcement activities, including patrol strategies, resource deployment, investigative priorities, and community engagement initiatives. By identifying crime hotspots, trends, and patterns, agencies can optimize patrol routes, allocate personnel to high risk areas, and enhance visibility in areas prone to criminal activities. Furthermore, crime risk assessment guides investigative efforts by providing insights into potential suspects, criminal networks, and patterns of criminal behaviour, facilitating more effective and efficient investigations [44].

Crime risk assessment also underscores the importance of integrating qualitative and quantitative data to create a comprehensive understanding of crime dynamics. While quantitative data, such as crime statistics and demographic information, provide measurable insights, qualitative data, including community feedback and anecdotal evidence, offer context and depth. The synthesis of these data types ensures that risk assessments are not only accurate but also reflective of the lived experiences of affected communities. This holistic approach enhances the relevance and applicability of crime prevention strategies, fostering trust and collaboration between law enforcement agencies and the communities they serve. Despite its numerous advantages, the implementation of crime risk assessment faces several challenges. Limited access to accurate and timely data, resource constraints, and gaps in expertise among Investigation Officers (IOs) can impede the effectiveness of risk assessment practices. Many IOs may rely on traditional approaches or personal judgment due to a lack of

specialized training in advanced methodologies. This reliance can lead to inconsistencies and inaccuracies in assessments, hindering the ability to identify and prioritize crime risks effectively. Addressing these challenges requires investment in capacity-building initiatives, including training programs, technological advancements, and policy reforms, to equip law enforcement personnel with the skills and tools necessary for effective risk assessment. Another critical aspect of crime risk assessment is its role in fostering community oriented policing. By engaging with community members and incorporating their perspectives into risk assessments, law enforcement agencies can develop more inclusive and contextually relevant strategies. Community input provides valuable insights into local crime concerns, social dynamics, and environmental factors that may influence criminal behaviour. This collaborative approach not only enhances the accuracy of risk assessments but also strengthens the relationship between law enforcement and the communities they serve. In addition to its operational benefits, crime risk assessment supports broader policy and strategic objectives within the criminal justice system. By providing evidence-based insights, risk assessments inform the development of policies aimed at reducing crime and enhancing public safety. For instance, data from risk assessments can guide the allocation of funding for crime prevention programs, infrastructure development, and technological innovations. Policymakers can leverage these insights to address systemic issues, such as socio economic disparities and urban planning challenges that contribute to crime. Technological advancements have further revolutionized crime risk assessment, enabling the integration of sophisticated tools and methodologies into law enforcement practices.



Figure 2.2 Illustration of Crime Risk Assessment Integrating Data Analytics, Community Input, and Strategic Policing to Identify Risks and Guide Preventive Actions

2.3 Global Crime Risk Assessment Systems

Crime risk assessment models are fundamental tools in modern law enforcement strategies, enabling the identification, evaluation, and mitigation of potential criminal activities. These models serve as the backbone of predictive and preventive policing, providing actionable insights that empower agencies to allocate resources efficiently, prioritize interventions, and address emerging crime trends. This section presents a precise and high-level examination of various existing crime risk assessment models, offering a detailed analysis of their methodologies, applications, and effectiveness while critically evaluating their limitations and areas for improvement. Over the past few decades, numerous crime risk assessment systems have been developed and implemented globally to support evidence-based decision making within the criminal justice process. These systems aim to evaluate the likelihood of reoffending, identify risk levels of offenders, and support strategic deployment of law enforcement resources. They incorporate a range of methodologies, including actuarial models, structured professional judgment, and machine learning algorithms, and are used at various stages including pretrial assessments, sentencing, parole decisions, and community supervision. One of the most widely recognized tools is the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS), developed in the United

States. COMPAS utilizes an actuarial approach to assess the risk of recidivism and violent behavior by analyzing static and dynamic offender characteristics. Despite its widespread adoption in jurisdictions such as New York, California, and Florida, COMPAS has faced criticism for alleged racial bias and a lack of transparency in its algorithmic decision making [45]. In the United Kingdom, the Offender Assessment System (OASys) is employed by Her Majesty's Prison and Probation Service to assess the risk and needs of individuals in custody or under community supervision. OASys combines structured professional judgment with actuarial data to inform rehabilitation and risk management strategies [46]. Another internationally utilized tool is Static-99R, a revised version of the Static-99, which is used to estimate the risk of sexual offender recidivism. This tool is actuarial in nature and incorporates historical and demographic variables such as age at release, victim profiles, and previous convictions. It has been validated across multiple jurisdictions including Canada, the United Kingdom, and Australia [47]. Canada has also pioneered the development of the Level of Service/Case Management Inventory (LS/CMI), which integrates risk assessment with intervention planning. This system evaluates offenders based on the principles of risk, need, and responsivity (RNR) and is widely implemented in correctional settings internationally [48]. Technological advancements have enabled the development of machine learning-based systems such as PredPol (Predictive Policing), which was initially piloted in Los Angeles. PredPol analyzes historical crime data to predict the time and location of potential future crimes, primarily focusing on property and drug-related offenses. While it has been lauded for its potential to enhance patrol efficiency, concerns have been raised regarding the reinforcement of biased policing practices due to historical data limitations [49]. In the United Kingdom, the Harm Assessment Risk Tool (HART) represents an innovative application of machine learning in policing. Developed by Durham Constabulary, HART uses a random forest algorithm trained on ten years of police data to classify individuals into high, moderate, or low risk categories. It has been employed in custody suites to support pretrial decision making and resource allocation [50,51]. For forensic and psychiatric populations, tools such as the Violence Risk Appraisal Guide (VRAG) are utilized to assess the risk of violent recidivism. VRAG is an actuarial instrument based on psychological assessments and criminal history, primarily used in mental health and correctional institutions in Canada. Youth specific tools have also emerged, such as the Structured Assessment of Violence Risk in Youth (SAVRY) and the Youth Level of Service/Case Management Inventory (YLS/CMI).

These instruments are tailored to adolescent populations and combine both structured professional judgment and actuarial elements to assess the risk of future violence and criminogenic needs [52,53]. These global systems demonstrate the evolution of crime risk assessment practices from purely judgment-based models toward more data-driven and algorithmic approaches. While each system varies in scope, methodology, and implementation context, they collectively highlight the importance of structured risk assessment in promoting public safety, informing policy decisions, and enhancing the efficiency of criminal justice operations.

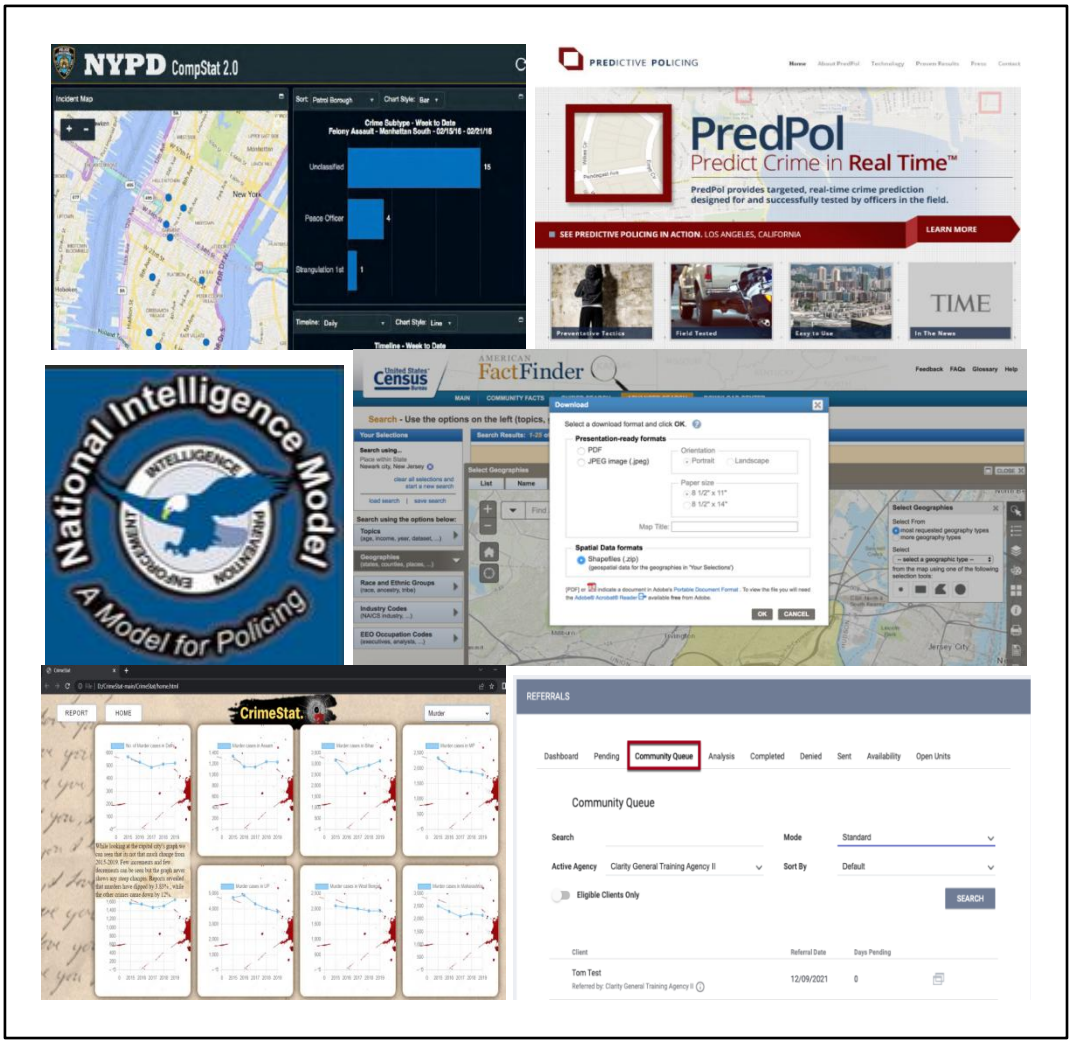


Figure 2.3 Visual Overview of Major Crime Risk Assessment Models Used Globally

The critique presented in this section will also delve into the benefits of these models, such as their ability to enhance decision making, optimize resource allocation, and improve operational efficiency. For example, models that incorporate advanced analytics and real-time data integration can provide law enforcement agencies with

timely insights, enabling swift and informed responses to emerging threats. Furthermore, the practicality of these models in diverse contexts will be explored, highlighting their adaptability to various policing strategies and community needs. At the same time, this critique will not shy away from identifying potential drawbacks and limitations inherent in these models. Common challenges include issues related to data quality, such as incomplete or outdated information, which can compromise the accuracy of predictions. The complexity of implementing certain models, particularly those requiring advanced technological infrastructure or specialized expertise, may also pose barriers to their widespread adoption. Additionally, the potential for biases in predictive algorithms, whether stemming from historical data or systemic factors, underscores the importance of ethical considerations in the development and deployment of these models. By offering this comprehensive analysis and critique, this section aims to provide a well rounded understanding of the current landscape of crime risk assessment models. This examination underscores their significance in shaping law enforcement strategies, highlighting their potential to transform how agencies approach crime prevention and intervention. At the same time, it emphasizes the need for on-going refinement and innovation to address existing limitations and adapt to the evolving challenges of crime and public safety.

Table 2.1
Summary and Critical Evaluation of Selected Crime Risk Assessment Systems: Country of Use, Target Population, Model Type, Purpose, and Key Benefits and Limitations

System	County	Target	Model Type	Purpose	Critique (Benefits & Limitations)
COMPASS	US	Adults	Actuarial	Recidivism, sentencing	Enhance sentencing decisions; criticized for racial bias and lack of transparency.
OASys	UK	Adults	Mixed	Probation, prison risk	Supports rehabilitation; implementation is resource-intensive and may rely on subjective inputs.
Static-99	Global	Adults / Sex offender	Actuarial	Recidivism prediction	Highly validated; limited to static factors, cannot reflect change over time.

System	County	Target	Model Type	Purpose	Critique (Benefits & Limitations)
LS/CMI	Canada / Global	Adults	Actuarial / Needs Based	General / violent risk	Integrates risk with case management; requires trained personnel and regular updating.
PredPol	US	Area	Machine Learning	Predictive policing	Offers real-time insights for patrols; criticized for reinforcing existing policing biases.
HART	UK	Individuals	Machine Learning	Pretrial harm risk	Uses big data for custody decisions; concerns over algorithmic fairness and transparency.
VRAG	Canada	Forensic Offenders	Actuarial	Violence prediction	Robust for psychiatric assessments; outdated data sources may reduce accuracy.
SAVRY	Global	Youth	Structured Clinical	Youth violence risk	Promotes individualized risk assessment; depends on clinician judgment, potential inconsistency.

2.4 Integrated Crime Risk Assessment Systems (ICRAS – Royal Malaysian Police)

In the evolution of technology, the rapid advancement of crime risk assessment tools has significantly transformed law enforcement practices across the globe. Leveraging innovations such as predictive analytics, machine learning, and real time data integration, modern policing has shifted from reactive to proactive strategies in crime prevention. This technological progression has empowered agencies to process vast volumes of data, uncover complex crime patterns, and derive actionable intelligence with unprecedented speed and accuracy. Moreover, enhanced data visualization platforms and inter agency data sharing mechanisms have facilitated greater collaboration, enabling coordinated responses across jurisdictions [54]. Recognizing these global developments and in preparation for the demands of the Fourth Industrial Revolution (IR 4.0), Malaysia is committed to not being left behind.

As part of its strategic vision for a modern and intelligent policing framework, the Integrated Crime Risk Assessment System (ICRAS) has been developed as a proactive and technology driven solution to enhance crime prevention efforts and law enforcement effectiveness nationwide. ICRAS integrates multiple data sources, analytical techniques, and technological tools to generate comprehensive insights into crime trends, risk factors, and vulnerable zones. Utilizing advanced algorithms and machine learning models, the system is capable of identifying high-risk individuals, locations, and behaviours allowing law enforcement agencies to optimize resource deployment and execute targeted interventions. This marks a significant departure from traditional reactive approaches, positioning ICRAS as a forward-looking initiative that aims to mitigate crime risks before they materialize. In alignment with international best practices, the implementation of ICRAS reflects Malaysia's progressive stance in adopting data-driven methodologies to strengthen public safety. As the global policing landscape continues to embrace innovation, systems like ICRAS will play a crucial role in shaping the future of crime prevention. However, the responsible deployment of such technologies must be accompanied by continuous capacity building, ethical oversight, and robust data governance to ensure effectiveness, fairness, and public trust.

2.5 Introduction to Crime Risk Assessment Models

Crime risk assessment models serve as structured frameworks designed to systematically evaluate and quantify the likelihood of criminal activities occurring within specific contexts or jurisdictions. These models integrate a vast array of data sources, analytical methodologies, and predictive algorithms to generate risk assessments, facilitating law enforcement agencies in resource allocation, intervention prioritization, and preventive measures. However, the efficacy and reliability of these models vary significantly based on their fundamental methodologies, data inputs, and validation procedures.

i. Actuarial Models

Actuarial models rely on statistical analyses of historical crime data to identify patterns and trends, subsequently predicting future criminal activities. Examples encompass the Risk-Need-Responsivity (RNR) model and the Violence Risk Appraisal

Guide (VRAG) [55]. These models have become pivotal in the field of criminology and predictive policing due to their capacity to process vast datasets and deliver objective, data-driven assessments. The RNR model, for instance, is designed to evaluate offenders' risks, needs, and responsivity to intervention strategies, offering a structured framework to reduce the likelihood of regression. Meanwhile, the VRAG employs psychometric tools and historical behavioural data to assess an individual's tendency for violent behaviour, enabling targeted preventative measures. Despite their strengths, actuarial models have been criticized for their inherent limitations. One major concern is their inability to account for contextual factors that influence criminal behaviour. These models often operate on predefined variables, such as past criminal records and socio-economic status, while neglecting dynamic elements like community support systems, cultural influences, and environmental stressors. Such failures can lead to oversimplified conclusions, reducing the effectiveness of interventions in diverse and complex social landscapes. Additionally, actuarial models tend to generalize individuals based on statistical averages, thereby disregarding the unique circumstances and personal motivations that drive criminal actions. For example, [56] argue that treating offenders as homogenous groups may lead to interventions that are ill-suited for specific cases. This approach undermines the principles of personalized justice and risks perpetuating systemic biases. Historical data, often used as the foundation for these models, may reflect existing disparities in law enforcement practices, such as over policing in marginalized communities. Consequently, predictions derived from such datasets can reinforce inequities, as highlighted by [57]. This has led to growing concerns about fairness, transparency, and ethical accountability in the deployment of actuarial models. Nonetheless, the objective and quantitative nature of actuarial models remains a significant advantage, particularly in resource limited settings. [58] emphasize that these models outperform unstructured clinical judgments in predicting criminal behaviour, especially among high risk populations. By systematically analysing data, they enable law enforcement agencies to allocate resources more efficiently, prioritize interventions, and implement preventative measures in areas with heightened risk. Such capabilities are indispensable in modern policing, where the demand for data driven strategies is increasingly critical. Recent advancements aim to address the limitations of traditional actuarial models by incorporating machine learning and artificial intelligence. These hybrid approaches can process larger datasets, identify more complex patterns, and adapt to contextual variables. [59] suggests that integrating

qualitative insights from experts and community stakeholders can further enhance the models' robustness, fostering more inclusive and equitable outcomes. In addition, while actuarial models like the RNR and VRAG offer valuable tools for predicting and mitigating criminal activities, their limitations necessitate careful consideration. Balancing their objective strengths with the need for contextual and individualized assessments is crucial for ethical and effective implementation. Continuous refinement, informed by interdisciplinary research and stakeholder engagement, holds the potential to transform actuarial models into more comprehensive frameworks for crime prevention and public safety.

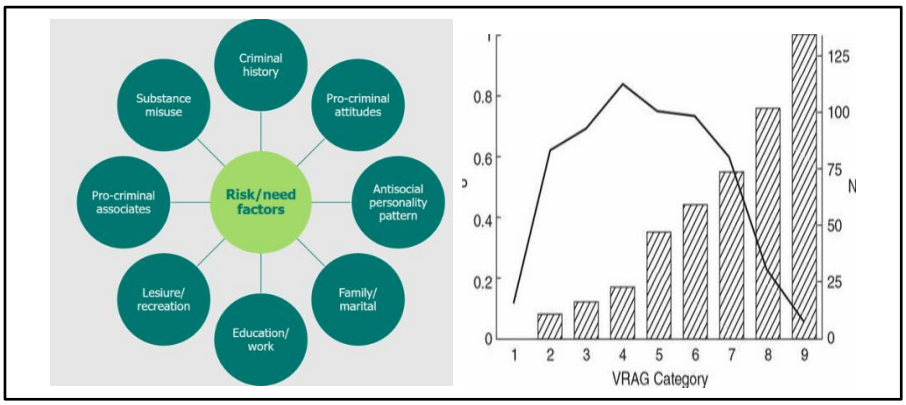


Figure 2.4 RNR and VRAG Approach in Actuarial Mode

ii. Clinical Models

Clinical models incorporate expert judgment and qualitative assessments to evaluate an individual's risk of engaging in criminal behaviour. These models often utilize structured interviews, psychological assessments, and case formulations to establish detailed and nuanced risk profiles. Examples include the Structured Assessment of Violence Risk in Youth (SAVRY) and the Historical Clinical Risk Management 20 (HCR-20) [60]. These tools emphasize a personalized approach to risk evaluation, integrating psychological insights with historical and contextual information to develop comprehensive assessments that go beyond statistical predictions. The SAVRY is specifically designed for assessing violence risk in adolescents and integrates dynamic risk factors such as peer influences, emotional regulation, and school adjustment with historical and clinical data. This model enables practitioners to tailor interventions to the unique developmental and environmental challenges faced by youth offenders [61]. Similarly, the HCR-20 combines historical

factors, such as past violent incidents, with clinical observations and situational elements to predict the likelihood of future violence. This model is widely recognized for its structured professional judgment approach, which balances empirical evidence with the nuanced expertise of trained evaluators [62]. One of the primary strengths of clinical models is their capacity to provide a holistic understanding of individual risk factors. Unlike actuarial models, which rely solely on statistical data, clinical models incorporate qualitative dimensions, such as personality traits, behavioural tendencies, and situational stressors. This multidimensional perspective allows for a deeper exploration of the underlying causes of criminal behaviour, facilitating more targeted and effective intervention strategies. For instance, case formulations derived from clinical assessments can identify specific triggers for violent behaviour, enabling tailored therapeutic or rehabilitative approaches [63].

Despite their strengths, clinical models are not without challenges. A key limitation is their susceptibility to biases and inconsistencies in expert judgment. The reliance on human evaluators introduces the potential for subjective interpretations, which may be influenced by personal beliefs, cultural perspectives, or even unconscious biases. Monahan (2006) highlights that inter-rater reliability in clinical assessments can vary significantly, raising concerns about the consistency and reproducibility of risk evaluations. Furthermore, research by [64,65] indicates that clinical models may overestimate or underestimate risk levels depending on the evaluator's experience and training, which can have profound implications for legal and policy decisions. Another limitation is the time intensive nature of clinical assessments. The process of conducting interviews, administering psychological tests, and formulating case profiles requires considerable resources and expertise, making these models less feasible in settings with high caseloads or limited access to trained professionals. Additionally, critics argue that the qualitative nature of clinical models makes them less transparent and harder to validate compared to actuarial approaches, which rely on quantifiable metrics and standardized algorithms [66]. To mitigate these challenges, recent advancements in clinical models have focused on enhancing standardization and integrating technology. Structured professional judgment tools, such as SAVRY and HCR-20, have introduced guidelines and scoring systems to improve consistency and reduce evaluator bias. Moreover, efforts to combine clinical assessments with data driven techniques, such as machine learning, are showing promise in bridging the gap between qualitative and quantitative approaches. For example, hybrid models that incorporate both clinical

judgment and actuarial data have demonstrated improved predictive accuracy and reliability [67].

SAVRY Structured Assessment of Violence Risk in Youth		Factor loading		
		1 Dissocial behavior	2 Assimilation	3 Maladaptation
Historical	History of violence	**0.530**	0.163	0.340
	History of non-violent offending	**0.647**	0.416	0.286
	Early initiation of violence	**0.367**	0.056	0.307
	Past supervision/intervention failures	**0.509**	0.133	0.202
	History of self-harm or suicide attempts	**0.286**	0.172	0.265
	Exposure to violence in the home	0.126	**0.544**	0.170
	Childhood history of maltreatment	0.116	**0.703**	0.215
	Parental/caregiver criminality	0.160	**0.666**	0.240
	Early caregiver disruption	0.262	**0.730**	0.152
	Poor school achievement	0.236	0.372	**0.406**
Social/contextual	Peer delinquency	**0.688**	0.276	0.097
	Peer rejection	0.035	0.150	**0.406**
	Stress and poor coping	0.104	0.242	**0.582**
	Poor parental management	0.273	**0.734**	0.206
	Lack of personal/social support	0.314	**0.609**	0.193
	Community disorganization	**0.482**	0.390	0.145
Individual	Negative attitudes	**0.638**	0.219	0.418
	Risk-taking/impulsivity	0.491	0.145	**0.535**
	Substance use difficulties	**0.606**	0.196	0.097
	Anger management problems	0.368	0.232	**0.623**
	Low empathy/remorse	0.265	0.141	**0.522**
	Attention-deficit/hyperactivity disorder	0.281	0.198	**0.488**
	Poor compliance	0.493	0.232	**0.493**
	Low interest/commitment to school or work	0.440	0.161	**0.434**

old values represent the factor loadings

Sub-Scales	Items
Historical Scale	
H1	Previous Violence
H2	Young Age at First Violent Incident
H3	Relationship Instability
H4	Employment Problems
H5	Substance Use Problems
H6	Major Mental Illness
H7	Psychopathy
H8	Early Maladjustment
H9	Personality Disorder
H10	Prior Supervision Failure
Clinical Scale	
C1	Lack of Insight
C2	Negative Attitudes
C3	Active Symptoms of Major Mental Illness
C4	Impulsivity
C5	Unresponsive to Treatment
Risk Management Scale	
R1	Plans Lack Feasibility
R2	Exposure to Destabilizers
R3	Lack of Personal Support
R4	Noncompliance with Remediation Attempts
R5	Stress

Figure 2.5 SAVRY and HCR-20

iii. Machine Learning Models

Machine learning has become a critical component in modern crime analysis, offering powerful tools to uncover behavioural patterns hidden within large and complex datasets. These models draw on computational techniques capable of processing vast volumes of structured and unstructured data to identify indicators of criminal activity. Unlike conventional statistical methods, which often rely on predefined assumptions, machine learning systems can detect intricate and previously unseen relationships, continually refining their predictive accuracy as new data become available. This adaptive quality has made them central to contemporary policing and crime prevention efforts. Among the models frequently discussed in recent literature are Neural Networks, Decision Trees, Support Vector Machines (SVMs), and Random Forests [68]. Neural Networks, inspired by the interconnected nature of neurons in the human brain, are particularly effective at recognizing nonlinear relationships between variables. Their layered architecture enables them to learn complex dependencies that can reveal subtle behavioural or spatial trends. Researchers have applied Neural Networks to tasks such as forecasting burglary hotspots, profiling offenders, and mapping criminal networks. For example, [69] demonstrated that Neural Network based models achieved notably higher accuracy in predicting burglary patterns compared with traditional regression based methods. Decision Trees, by contrast, are valued for their clarity and interpretability. They classify data through a sequence of conditional splits based on attributes such as crime type, location, and time of occurrence. Because each decision path can be visualized and explained, these models are often employed in investigative prioritization and fraud detection, where transparency is crucial [70]. Support Vector Machines (SVMs) represent another established approach within predictive crime analysis. They operate by finding the optimal boundary that separates categories within the data, a property that makes them particularly effective for anomaly detection and classification tasks.

SVMs have been successfully used to identify suspicious financial transactions and other rare events within large datasets, as noted by [71]. Meanwhile, the Random Forest algorithm extends the decision tree concept by combining multiple trees to form a more stable and accurate ensemble model. Through random sampling of features and data subsets, Random Forests mitigate the risk of over fitting and improve generalization. They have been applied to a range of criminal intelligence problems,

including risk assessment, hotspot mapping, and recidivism prediction. Studies such as [72] report that Random Forests often outperform single decision trees, owing to their ability to capture more nuanced relationships across variables while maintaining a degree of interpretability. Even with these advances, the application of machine learning in criminology is not without controversy. One persistent challenge lies in the bias embedded within historical data. When models are trained on data reflecting systemic inequalities such as disproportionate arrest rates among particular demographic groups they risk reinforcing those very disparities. [73] highlighted this issue by showing how racial bias present in historical records influenced the predictions of recidivism risk models. Such findings have driven a growing interest in fairness-aware algorithms and responsible AI practices aimed at ensuring equity in predictive outcomes [74]. Another important consideration concerns the transparency of these systems. Complex models, particularly deep learning architectures, are often criticized as "black boxes" because they provide little explanation for how a given output is derived. This lack of interpretability poses challenges for accountability in policing and judicial contexts, where decisions must be justified.

Recent efforts to address this limitation have focused on explainable AI (XAI) frameworks such as LIME and SHAP [75–77], which aim to clarify how specific variables contribute to a model's prediction. While these tools improve interpretability, they also demand substantial computational resources and technical expertise. Moreover, inconsistent or incomplete crime data still common in many jurisdictions—can further constrain model reliability. To mitigate these gaps, researchers have begun integrating auxiliary data sources such as social media activity, IoT sensor feeds, and geographic information systems (GIS). These complementary datasets enhance situational awareness and strengthen predictive models [78]. In practice, machine learning has already shown value in predictive policing, guiding patrol allocation, and detecting emergent crime hotspots. Natural Language Processing (NLP) methods have further extended these capabilities, enabling real-time analysis of police narratives, online communications, and public sentiment [79]. Together, these approaches reflect how machine learning through models like neural networks, SVMs, decision trees, and Random Forests has reshaped the analytical toolkit of modern law enforcement. Yet, as the technology continues to advance, questions surrounding fairness, interpretability, and ethical governance remain as central to the field as the pursuit of accuracy itself.

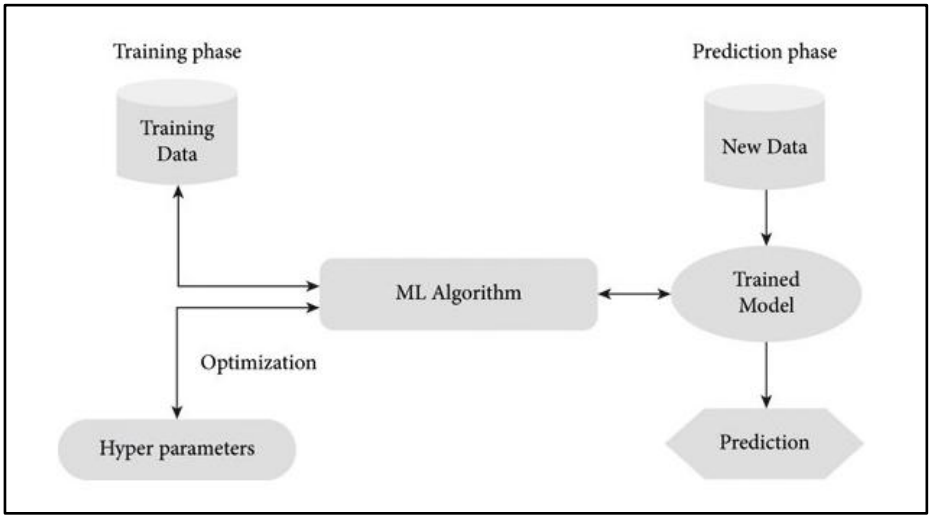


Figure 2.6 General Machine Learning Overview

2.5.1 Critical Evaluation of Existing Crime Risk Assessment Models

The development of crime risk assessment has been shaped by various predictive models and analytical tools designed to assist law enforcement agencies in identifying crime patterns and managing offender risks. These systems have contributed valuable insights to criminological research and operational policing. However, despite their contributions, many existing models exhibit methodological, ethical, and operational limitations that restrict their effectiveness in complex policing environments. One widely discussed predictive policing system is PredPol, which forecasts crime hotspots based on historical crime data and geospatial patterns. Although the system has demonstrated some ability to identify general crime trends, its heavy reliance on historical data has raised concerns regarding potential bias. Critics argue that such systems may reinforce existing policing patterns by repeatedly directing enforcement toward locations that historically recorded higher crime rates. In addition, PredPol has been criticized for its limited transparency, as its predictive mechanisms operate largely as a "black-box," making it difficult for practitioners to understand how predictions are generated. Similarly, the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) system, commonly used to estimate offender recidivism risk, has generated debate regarding fairness and algorithmic bias. Several studies suggest that COMPAS may disproportionately classify minority defendants as high-risk, raising concerns about the ethical implications of algorithm-based decision support tools in criminal justice. The limited transparency of the algorithm further complicates the ability of legal professionals to verify or interpret its risk

assessments.The Offender Assessment System (OASys), used in the United Kingdom, provides a more comprehensive evaluation by incorporating behavioural and social factors in offender risk assessments. However, the system still relies significantly on static indicators such as criminal history and past behaviour. This reliance may limit its ability to capture dynamic changes in offender behaviour or evolving environmental conditions. Additionally, the absence of adaptive learning mechanisms reduces its capacity to respond to emerging crime trends.

Other risk assessment tools, including Static-99R and the Level of Service/Case Management Inventory (LS/CMI), also depend heavily on historical and fixed risk indicators. While these factors provide baseline insights, they often fail to reflect the dynamic and contextual nature of criminal behaviour. As a result, these models may have limited effectiveness when applied to complex real-world policing environments.The Harm Assessment Risk Tool (HART) represents a more recent attempt to incorporate machine learning into crime risk assessment. Although HART introduces advanced analytical capabilities, concerns remain regarding potential bias and the influence of socio-economic variables in prediction outcomes. Furthermore, the model may lack flexibility when applied across different jurisdictions with varying crime patterns.Overall, while these models have contributed significantly to the development of crime risk prediction, several limitations remain. Many systems rely heavily on historical data and static risk indicators, offer limited algorithmic transparency, and raise ethical concerns related to bias and fairness. In addition, most existing tools lack the capability to integrate real-time data or continuously adapt to evolving crime patterns.

To address these limitations, the Integrated Crime Risk Assessment System (ICRAS) was developed as a more comprehensive and adaptable framework. Unlike many existing models that rely primarily on historical data, ICRAS integrates multiple data sources, including crime incident records, socio-economic indicators, and geospatial information. This integration enables the system to capture complex interactions among factors influencing crime patterns.ICRAS employs the Random Forest machine learning algorithm to support predictive analysis. Random Forest is well suited for modelling complex and non-linear relationships within large datasets while reducing the risk of overfitting. This approach enables the system to generate stable predictions and improve its performance as new data becomes available. In addition, ICRAS emphasizes interpretability and operational usability by combining predictive

analytics with visualisation tools that provide clear and actionable insights for investigation officers. This design improves transparency and supports more informed decision-making compared to several existing predictive policing systems. Furthermore, ICRAS represents a shift from reactive crime response toward proactive crime prevention. By integrating diverse data sources with machine learning-based predictive modelling, the system provides a more robust framework for identifying emerging crime risks and supporting evidence-based policing strategies.

2.5.2 Analysis of Strengths and Weaknesses

a) Strengths

Crime risk assessment models furnish law enforcement agencies with systematic tools for identifying and prioritizing crime prevention efforts, efficiently allocating resources, and effectively targeting interventions. These models provide a structured and standardized approach to assessing risk, enabling a transition from reactive to proactive crime management strategies. By leveraging these models, law enforcement agencies can identify high-risk areas, vulnerable populations, and emerging trends, thereby focusing their efforts on areas with the greatest potential impact. This systematic methodology not only enhances operational efficiency but also fosters accountability by basing decisions on quantifiable data and predictive analytics. These models promote resource optimization, a crucial consideration given the budgetary and personnel constraints faced by many law enforcement agencies. By pinpointing areas with the highest likelihood of criminal activity, agencies can allocate patrols, investigative teams, and community engagement efforts more effectively. For example, [80] highlighted how predictive policing models enabled police departments in urban areas to reduce response times and improve crime resolution rates by focusing on priority zones. Another advantage is the facilitation of data-driven decision-making. Traditional methods often rely on the intuition or experience of individual officers, which, while valuable, can introduce subjectivity and inconsistencies. In contrast, risk assessment models use evidence based approaches that minimize biases and enhance the reliability of decision-making processes. [81] emphasized that the use of machine learning algorithms in risk assessment has significantly improved the accuracy of predictions, particularly for complex crimes such as financial fraud and organized

crime. Collaboration and information sharing among stakeholders are also bolstered by these models. Standardized frameworks for risk assessment provide a common language and methodology, enabling smoother communication between law enforcement, government agencies, community organizations, and academic researchers. This collaborative environment fosters innovation, as evidenced by partnerships between police departments and data scientists to develop cutting edge analytical tools.

Moreover, these models support transparency and accountability in law enforcement operations. By relying on data and standardized methodologies, they provide clear justifications for resource allocation and intervention strategies. This transparency is critical for maintaining public trust, particularly in communities where law enforcement practices have historically faced scrutiny. [82] argued that the adoption of fairness aware algorithms within risk assessment models has the potential to mitigate biases and promote equitable outcomes across diverse populations. Furthermore, crime risk assessment models also offer a strong framework for modernizing and optimizing law enforcement strategies. Their strengths lie in their systematic and evidence-based approach, ability to integrate diverse data sources, resource optimization, promotion of collaboration, adaptability, and enhancement of transparency and accountability. As technological advancements continue to refine these models, their potential to transform public safety and crime prevention efforts remains immense. However, a balanced analysis requires an examination of their limitations, which will be discussed in subsequent sections.

b) Weaknesses

Despite their significant contributions, crime risk assessment models are not without limitations. Actuarial models, while grounded in statistical rigor, often fail to account for the broader socio-cultural and contextual nuances that influence criminal behaviour. Their reliance on historical data can inadvertently reinforce existing biases, particularly in cases where past data reflect systemic inequalities or inaccuracies. Furthermore, these models tend to generalize risk across populations, which may obscure critical individual or situational factors essential for precise assessments [83]. Similarly, clinical models, which rely on professional judgment and expertise, introduce inherent subjectivity and variability into the risk assessment process. The reliance on

individual discretion, while valuable for its adaptability, can lead to inconsistent evaluations across different practitioners or jurisdictions. This variability undermines the standardization and comparability of results, limiting the model's reliability in broader applications [84]. Machine learning models are impressive for their ability to handle massive amounts of data and find complex patterns, but they also come with their own set of challenges. These models are heavily influenced by the quality and representativeness of their training data. If the data used is biased or incomplete, the model's predictions may perpetuate or even exacerbate these biases, leading to skewed outcomes that disproportionately affect certain communities. Additionally, the lack of interpretability or transparency in many machine learning algorithms often referred to as the "black box" problem, hampers trust and accountability. Law enforcement agencies and policymakers may struggle to understand or justify the model's predictions, raising ethical and operational concerns [85].

Another critical limitation across all models is the potential for out dated or incomplete data to skew results. Crime risk assessment models require robust, up to date datasets to remain effective. However, in many cases, data collection processes are hindered by bureaucratic constraints, privacy concerns, or technological limitations. These gaps can result in assessments that do not accurately reflect current crime dynamics, reducing their utility for proactive policing and prevention strategies [86]. Ultimately, while crime risk assessment models provide invaluable tools for law enforcement, addressing these weaknesses is imperative to ensure their efficacy, fairness, and ethical application. By recognizing and mitigating these limitations, practitioners can develop more robust and inclusive systems that balance predictive accuracy with contextual sensitivity.

Table 2.2
Strengths and Weaknesses of Crime Risk Assessment Models in Law Enforcement

Category	Strength	Weaknesses
Analytical Rigor	- Enables systematic, standardized risk analysis. - Transitions law enforcement from reactive to proactive strategies.	- Actuarial models often fail to capture socio-cultural contexts. - Overgeneralization across populations can overlook individual or situational nuances.

Category	Strength	Weaknesses
Data Driven Decision-Making	- Supports evidence-based practices. - Reduces subjectivity compared to traditional methods.	- Machine learning models rely heavily on the quality and representativeness of training data. - Biased data can lead to skewed predictions.
Resource Optimization	- Enhances allocation of personnel and budgets based on risk prioritization.	- Inaccurate or outdated data can compromise resource deployment effectiveness.
Operational Efficiency	- - Improves response time and resolution rates (e.g., predictive policing in urban areas).	- Bureaucratic or technical issues may limit access to timely, high-quality data.
Collaboration and Integration	- Promotes inter-agency communication via standardized frameworks. - Enables partnerships with academia and tech experts.	- Variability in clinical models undermines consistency and comparability.
Transparency and Accountability	- Data-backed justifications for strategies build public trust. - Fairness-aware algorithms promote equity.	- - Machine learning "black box" models lack interpretability, challenging oversight and public accountability.
Scalability and Innovation	- Flexible integration with technological tools and real-time analytics. - Adaptive to evolving crime trends.	- Implementation may require specialized expertise and advanced infrastructure, limiting use in under-resourced jurisdictions.

2.5.3 Case Studies and Practical Applications

The effective development and deployment of crime risk assessment models like ICRAS require a detailed understanding of real world scenarios and lessons learned from practical applications. Case studies offer invaluable insights into the successes, challenges, and strategies employed by law enforcement agencies to implement such models. These studies serve as evidence-based references that guide the adoption of best practices while highlighting potential pitfalls that may arise during deployment. One of the keys focuses of case studies is the exploration of implementation challenges**,** which can vary significantly across different jurisdictions. Law enforcement agencies often encounter issues such as inadequate data quality and availability, resource constraints, and organizational resistance. For instance, predictive models implemented in resource limited environments, such as certain rural or developing regions, may face hurdles in acquiring comprehensive and consistent crime data. A study by [87] illustrated how resource limitations in smaller municipalities hindered the long term

effectiveness of predictive policing systems, emphasizing the importance of scalable and cost-effective solutions. Another critical area of focus is community engagement**,** which is essential for fostering trust, transparency, and accountability. Case studies of community-driven initiatives have shown that participatory approaches enhance the cultural sensitivity and inclusivity of crime risk models. In Chicago, for example, feedback from community organizations led to adjustments in predictive models to address concerns about racial profiling and over policing [88]. By involving community representatives in oversight committees and co design processes, law enforcement agencies can align their tools with public expectations and build stronger partnerships with the communities they serve. By analysing real-world examples, law enforcement agencies can adopt evidence based strategies that enhance the effectiveness, fairness, and adaptability of crime prevention systems.

a) Implementation Challenges: Real world implementation of crime risk assessment models often faces various challenges, including data quality issues, resource constraints, and organizational resistance. Case studies illustrating successful implementation strategies and lessons learned can provide valuable insights for law enforcement agencies. These challenges span technical, resource related, and organizational dimensions, requiring a strategic and systematic approach to address them. A critical challenge lies in the quality and availability of data, which forms the foundation for predictive accuracy. Predictive models depend on extensive datasets, including historical crime records and demographic information, to identify patterns and generate actionable insights. However, these datasets are often incomplete, inconsistent, or out dated, which can undermine the reliability of predictions and decision making. For example, data inconsistencies have been reported in certain regions where limited technological infrastructure or fragmented record keeping practices result in incomplete historical crime records. Additionally, biases embedded in historical data can distort predictive outcomes, as seen in cases where crime data disproportionately represented over policed communities [89]. To address this, agencies must establish rigorous data collection, validation, and integration protocols, ensuring that the datasets used are accurate, comprehensive, and reflective of current realities. Another pressing issue is

resource constraints, which pose significant barriers to the adoption and maintenance of advanced crime risk assessment systems. Developing and deploying these systems require substantial investment in technology, infrastructure, and specialized personnel. Law enforcement agencies operating under limited budgets often struggle to secure funding for such initiatives, particularly in smaller jurisdictions or regions with resource-strapped public safety departments. Additionally, the continuous operation of these systems requires periodic software updates, system maintenance, and on-going training of personnel, further straining limited resources. For instance, agencies with constrained budgets may face delays in implementing critical updates or may be unable to retain skilled professionals, thereby compromising the system's long-term effectiveness [90]. Securing sustained funding and exploring collaborative partnerships with external organizations can help alleviate some of these constraints. Organizational resistance remains another formidable obstacle to successful implementation. Within law enforcement agencies, scepticism toward new technologies often stems from concerns about disrupting established practices or replacing professional judgment with algorithmic decision-making. Officers may resist relying on predictive tools if they perceive them as undermining their expertise or introducing complexity into operational workflows. For example, early adopters of predictive policing systems in some metropolitan areas reported significant resistance from frontline officers, who questioned the reliability and practicality of algorithm driven recommendations [91]. To overcome this, it is essential to foster organizational buy-in through inclusive approaches. Engaging officers during the design and testing phases, demonstrating how the system can enhance rather than replace traditional practices, and providing comprehensive training programs can build trust and acceptance. Moreover, leadership support and clear communication of the system's benefits are vital in aligning agency wide goals with the new technology. Finally, challenges related to public perception and trust cannot be overlooked. Public support is critical for the successful deployment of crime risk assessment systems, as communities directly experience their outcomes. However, advanced technologies can evoke concerns about misuse or lack of transparency, particularly in communities with strained relationships with law enforcement. For example, in several U.S. cities such as in Chicago, Los

Angeles and New Orleans, predictive policing initiatives faced public opposition due to fears of increased surveillance and perceived targeting of specific communities. To address these concerns, agencies must prioritize community engagement by involving local stakeholders in discussions about the system's goals and benefits. Transparent communication, participatory approaches, and visible safeguards to prevent misuse are essential in building trust and ensuring community cooperation. The implementation of crime risk assessment models such as ICRAS involves addressing significant challenges related to data quality, resource allocation, organizational resistance, and public perception. Overcoming these barriers requires strategic planning, robust training, and proactive stakeholder engagement to ensure that these systems are effectively integrated and sustained in law enforcement operations.

b) Community Engagement: Engaging communities in the development and implementation of crime risk assessment models fosters trust, transparency, and accountability. Case studies highlighting community driven initiatives and participatory approaches can inform best practices for inclusive and culturally sensitive risk assessment processes. Community involvement ensures that the model reflects the unique characteristics, priorities, and concerns of the population it serves, thereby increasing its acceptance and effectiveness in addressing localized safety challenges. By actively involving community members during the design and development phases, law enforcement agencies can incorporate valuable insights that enhance the model's relevance and applicability. For instance, participatory workshops or focus groups can help identify specific risk factors, crime hotspots, and social dynamics that may not be readily apparent through traditional data sources. These collaborative efforts not only improve the predictive accuracy of the model but also empower residents to take an active role in shaping public safety strategies, fostering a sense of ownership and shared responsibility. Community-driven initiatives have demonstrated their effectiveness in diverse socio-economic and cultural contexts. For example, projects in cities like Oakland, California, have emphasized collaborative partnerships between law enforcement and community organizations to co-create solutions for crime prevention. By integrating local knowledge and cultural sensitivities, these initiatives have led to more equitable and effective resource allocation, addressing both immediate

safety concerns and underlying social issues [92]. Transparency is another critical element of successful community engagement. Openly communicating the objectives, methods, and limitations of crime risk assessment models builds trust and dispels misconceptions about their purpose. For example, hosting public forums or releasing regular updates about the system's performance can reassure communities that the model is being used responsibly and equitably. Moreover, involving community representatives in oversight mechanisms, such as advisory boards, ensures that decisions made using these models are aligned with public expectations and values. Case studies also highlight the importance of participatory approaches in mitigating potential biases and ensuring inclusivity. In Chicago, for example, community driven feedback on predictive policing tools helped to address concerns about racial disparities in crime data and its impact on risk assessments. This collaboration led to adjustments in the model's algorithms and the development of safeguards to prevent over policing in vulnerable neighbourhoods [93]. Such examples underscore the value of community input in creating fair and culturally sensitive systems. Education and awareness campaigns further strengthen community engagement by demystifying crime risk assessment models and fostering informed public participation. These campaigns can explain how the models work, the types of data they use, and the safeguards in place to protect against misuse. In jurisdictions where public hesitation about technology in law enforcement is high, such initiatives have proven essential for building support. For example, in several U.S. cities, outreach programs that clarified the goals and benefits of predictive policing technologies significantly improved public perception and cooperation [94]. Moreover, engaging communities in the development and implementation of crime risk assessment models is fundamental to their success. By fostering trust, promoting transparency, and incorporating diverse perspectives, community engagement enhances the social legitimacy and operational effectiveness of these systems. Best practices derived from case studies of participatory initiatives provide valuable guidance for law enforcement agencies seeking to implement inclusive and culturally sensitive crime risk assessment processes.

2.6 Predictive Modelling

Predictive modelling has emerged as one of the most transformative tools in modern law enforcement, providing law enforcement agencies with the ability to forecast crime, allocate resources efficiently, and implement proactive policing strategies. Unlike traditional methods that primarily rely on historical data to respond to crime after it occurs, predictive modelling leverages both historical and real time data to anticipate potential criminal activities. Through the application of sophisticated statistical techniques and machine learning algorithms, predictive models identify patterns, trends, and risk factors that can inform decision making, ultimately leading to more effective crime prevention and intervention [95]. This paradigm shifts from reactive to proactive policing is crucial, especially in urban environments where resources are limited, and the demand for timely, data-driven interventions is high [96]. At its core, predictive modelling in law enforcement seeks to predict the "when," "where," and "who" of criminal activity by analysing large datasets that include historical crime reports, socio-economic factors, demographic information, and environmental variables. By processing these datasets, predictive models are able to highlight crime hotspots, forecast trends in criminal behavior, and even identify individuals at risk of becoming involved in criminal activities [97]. These insights enable law enforcement agencies to allocate personnel and resources to areas with the highest likelihood of criminal occurrences, thus reducing the occurrence of crimes and enhancing overall public safety [98]. The advent of machine learning in predictive policing further enhances the capabilities of these models by introducing dynamic, self learning algorithms that continuously improve their accuracy over time. Techniques such as decision trees, random forests, and neural networks have been employed to uncover hidden patterns in large datasets, increasing the precision of crime predictions. The use of machine learning algorithms has been instrumental in real time crime forecasting, where patterns emerge quickly, and predictive models must adapt to new data [99]. However, as machine learning models become more complex, concerns about transparency and accountability in decision making processes have arisen. There are fears that these models might perpetuate biases if not carefully monitored, reinforcing systemic inequalities in law enforcement practices [100]. Therefore, it is essential for law enforcement agencies to adopt ethical frameworks that guide the development, implementation, and evaluation of predictive models.

While predictive modelling offers significant benefits, it is not without its challenges. A major concern is the quality and completeness of the data used in predictive models. Crime data often contains biases, such as over policing in certain communities or gaps in reporting, which can skew predictions and lead to disproportionate targeting of specific groups. To address these issues, it is critical to ensure that the data used for predictive modelling is comprehensive, representative, and regularly updated. Furthermore, integrating community feedback into the development and deployment of predictive models can help mitigate bias, foster trust between law enforcement and the public, and improve the accuracy of predictions [101]. The integration of real-time data streams also adds a layer of complexity to predictive modelling. As cities increasingly deploy surveillance systems, sensors, and social media monitoring tools, real-time data provides an unprecedented opportunity to refine crime predictions. By incorporating these dynamic data sources, predictive models can adapt to rapidly changing conditions, providing law enforcement with near-instant insights into emerging crime trends [102]. However, the use of real time data also raises questions about privacy, surveillance, and data security, which need to be carefully managed to maintain public trust. Predictive modelling has revolutionized law enforcement strategies by providing a data-driven approach to crime prevention. From hotspot analysis to machine learning, predictive models allow agencies to anticipate criminal activity and allocate resources more effectively. Despite its potential, the implementation of predictive modelling must be carefully managed to ensure that data biases, ethical concerns, and community trust are addressed. As predictive models continue to evolve, they offer the promise of more efficient, equitable, and proactive policing, contributing to safer communities and enhanced public safety. The adoption of predictive modelling in law enforcement has gained momentum in recent years, driven by advancements in technology, the availability of vast amounts of data, and the increasing demand for proactive crime prevention strategies. Law enforcement agencies across the globe have embraced predictive modelling as a means to enhance their capabilities in combating crime and ensuring public safety. From large metropolitan police departments to small-town sheriff's offices, predictive modelling has become an integral component of modern policing practices. The importance of predictive modelling in crime prevention cannot be overstated. Traditional reactive approaches to law enforcement, which rely primarily on responding to incidents after they occur, are inherently limited in their effectiveness. By contrast, predictive modelling enables law

enforcement agencies to adopt a proactive stance, anticipating and preventing criminal activities before they happen.

Predictive modelling ability to identify crime hotspots has been widely validated through real-world applications where law enforcement agencies have achieved significant crime reductions by strategically targeting high-risk areas, often utilizing tools such as Geographic Information Systems (GIS) and statistical algorithms to analyse historical crime data and forecast future incidents, as exemplified by the Los Angeles Police Department's implementation of the PredPol system, which analysed patterns of property crimes like burglary and vehicle theft to allocate police patrols to specific locations and times with high predicted crime risks, resulting in a measurable decline in these crimes compared to areas using traditional methods, a success largely attributed to the deterrent effect of an increased police presence in targeted zones [103]. Similarly, in the United Kingdom, Kent Police adopted predictive policing techniques that analysed historical crime data, including the time, location, and type of offenses, to forecast potential crime hotspots, enabling strategic resource deployment that significantly reduced incidents of burglary and vehicle theft, highlighting the operational advantages of integrating predictive analytics into routine policing strategies, while in Boston, Operation Ceasefire employed a hotspot-focused approach to address gang-related violence by identifying neighbourhoods with elevated risks of violent incidents, focusing both police enforcement and community outreach efforts in these areas, and achieving a 63% reduction in youth homicides in targeted locations, a result that underscores the potential of predictive modelling when combined with community oriented and multi agency approaches [104].

Additionally, the New York Police Department's CompStat program provides another compelling example, as it integrates real-time and historical crime data to generate dynamic heat maps and identify emerging crime patterns, which have enabled rapid responses to localized surges in violent crime and the strategic allocation of resources to mitigate these threats, demonstrating the scalability and adaptability of predictive models in diverse urban environments, while emerging applications in tackling cybercrime, such as the use of digital crime heat maps to identify areas with high incidents of online fraud or phishing, reflect the versatility of predictive analytics in addressing traditional and modern crime challenges alike. These systems typically rely on advanced algorithms, such as kernel density estimation (KDE) to create visual heat maps of high intensity crime areas, or machine learning classifiers that dynamically

adapt to new data, drawing from diverse data sources, including historical crime trends, temporal factors like seasonality and time of day patterns, and environmental and demographic information [105] , all of which contribute to the precision and relevance of hotspot predictions, though challenges persist in ensuring the ethical use of these technologies, as critics warn that biased historical data can lead to over-policing or discriminatory practices, necessitating measures such as community consultation, transparency initiatives, and regular audits to ensure fairness and accountability, as evidenced by the transparency focused practices adopted in Oakland, California, which emphasize the importance of aligning predictive systems with principles of equity and justice. Thus, predictive modelling, demonstrated through the successful outcomes of programs in Los Angeles, the UK, Boston, and New York, along with its potential in addressing emerging threats such as cybercrime, remains a transformative tool in modern law enforcement by enabling proactive, data-driven interventions that enhance operational efficiency and public safety, while on-going ethical oversight and community engagement are essential to maximizing its benefits and mitigating its risks.

Moreover, predictive modelling facilitates more efficient resource allocation, such as deploying officers to high crime areas during peak times, assigning specialized units to specific tasks like surveillance or community outreach, and strategically placing equipment such as surveillance cameras or mobile command units in locations identified as high risk. By enabling law enforcement agencies to optimize their limited resources and manpower, predictive models allow for the prioritization of interventions based on data-driven insights, ensuring that personnel, equipment, and funding are directed to areas and activities with the highest potential for crime prevention impact. For instance, predictive analytics can help determine when and where additional patrols are needed, reducing response times and increasing police visibility in vulnerable areas, or identify specific neighbourhoods requiring enhanced community policing efforts, fostering trust and collaboration between law enforcement and residents. This targeted approach enhances the overall effectiveness of law enforcement efforts, minimizes resource wastage, and maximizes the return on investment in crime prevention initiatives, ultimately contributing to safer communities and more sustainable policing practices.

In addition to its operational benefits, predictive modelling has profound implications for community policing and public trust. By proactively addressing crime hotspots and collaborating with communities to prevent criminal activities, law

enforcement agencies can foster positive relationships with residents, businesses, and other stakeholders. For example, the Chicago Police Department's Strategic Decision Support Centers (SDSCs) leverage predictive analytics to identify high-crime areas and organize community-based interventions. These centres work closely with community leaders to design tailored crime prevention strategies, such as organizing neighbourhood watch programs or hosting safety workshops, which empower residents to play an active role in crime prevention efforts. Similarly, in Santa Cruz, California, the use of predictive policing software has not only enabled targeted patrolling in high-risk areas but has also opened avenues for dialogue between officers and community members about safety concerns and the rationale behind resource allocation. Such proactive engagement enhances public confidence in law enforcement, fosters cooperation and collaboration, and ultimately contributes to safer and more resilient communities by ensuring that crime prevention is a shared responsibility rooted in mutual trust and accountability.

2.6.1 Predictive Modelling in Law Enforcement

Predictive modelling has revolutionized contemporary law enforcement by enabling agencies to anticipate and mitigate crime rather than merely responding to it after the fact. This transformation stems from the ability of predictive models to harness historical and real time data, analyse patterns, and uncover risk factors that contribute to criminal activities. These models utilize advanced algorithms to forecast crime hotspots, identify potential offenders, and allocate resources more effectively, thus embodying a proactive approach to crime prevention. [106] in their influential RAND Corporation report, examined how predictive policing models can be used to anticipate and prevent criminal activity. They highlighted several key methods and approaches that have proven effective in real-world applications. One prominent method discussed was hotspot analysis, which uses spatial and temporal crime data to identify areas where crimes are likely to occur. By applying techniques such as kernel density estimation (KDE) and nearest neighbour analysis, law enforcement agencies can map crime clusters and focus their patrols on high-risk locations. The study also explored the use of regression models to predict crime trends. These models analyse relationships between variables, such as time of day and environmental factors, to determine patterns that influence criminal behaviour. Another approach highlighted was Risk Terrain

Modelling (RTM), which links environmental characteristics like poorly lit areas or high traffic locations to the likelihood of crime. This method allows agencies to identify and reduce crime risks within specific locations. Although machine learning tools were less commonly applied at the time, Perry et al. acknowledged their growing potential for clustering and classification tasks. These tools can analyse large datasets to uncover patterns and forecast future criminal activity. The report emphasized the importance of diverse and high quality data inputs, including historical crime records and socioeconomic variables, for accurate predictions. Perry et al. supported their analysis with real-world case studies, such as the Los Angeles Predictive Policing Experiment (Operation LASER) and the Santa Cruz Predictive Policing Initiative. These examples demonstrated how predictive models helped reduce crime rates and optimize resource allocation. However, the study also highlighted limitations, such as data quality issues and the risk of perpetuating systemic biases through the reliance on historical data.

One of the most widely adopted techniques within predictive modelling is hotspot analysis, which focuses on identifying geographic areas with a high concentration of criminal activity. This method employs statistical tools such as kernel density estimation and spatial temporal clustering to pinpoint specific locations where crimes are likely to occur. Studies by [107] have demonstrated how this approach has significantly reduced crime rates in urban settings. For instance, in a landmark study in Los Angeles, hotspot analysis led to a 33% decrease in property crimes when law enforcement redirected their patrols to areas flagged as high-risk zones. This success is further amplified by the integration of Geographic Information Systems (GIS), which visualize these hotspots on dynamic maps, allowing officers to monitor real-time crime trends and respond swiftly to potential threats [108]. Predictive modelling also incorporates offender profiling, leveraging machine learning algorithms to predict the likelihood of individuals committing crimes based on behavioural patterns and historical records. This technique was notably employed in the Chicago Police Department's Strategic Subjects List (SSL), which aimed to identify individuals at risk of violent crimes [109]. By analysing factors such as prior arrests, age, and network associations, the SSL demonstrated a capacity to predict involvement in shootings with moderate accuracy. However, critics have pointed out that such models risk perpetuating systemic biases if underlying datasets reflect historical inequities in policing practices [110]. A notable example of predictive modelling success is the application of the Predictive Policing Initiative in Santa Cruz, California. This program

utilized historical burglary data combined with environmental variables, such as weather and time of day, to predict where property crimes were likely to occur. Law enforcement officers deployed to these predicted hotspots reported a measurable decrease in crime rates over a six-month period [111]. Similar efforts have been observed in New York City, where predictive models were applied to subway-related crimes, leading to enhanced security measures and reduced incidences of theft [112].

Despite these successes, real-world applications of predictive modelling raise critical methodological and ethical questions**.** For example, the reliance on large datasets introduces challenges related to data quality, completeness, and representativeness. Inaccurate or biased data can lead to flawed predictions that disproportionately target marginalized communities, exacerbating social inequalities [113]. To address this, some law enforcement agencies, such as those in Seattle and Boston, have adopted community driven approaches that integrate public feedback into their models. These participatory methods aim to enhance transparency, foster trust, and align predictive modelling practices with the needs and concerns of local populations. From an operational perspective, the deployment of predictive modelling necessitates a multi-faceted approach that combines technology, training, and collaboration**.** Agencies must invest in advanced infrastructure, such as cloud-based data storage and computational resources, to support the scalability of predictive systems. Additionally, training law enforcement personnel to interpret and act on model outputs is critical for bridging the gap between algorithmic predictions and on-ground enforcement actions [114]. Collaborative partnerships with academic institutions and private-sector experts have also proven valuable in refining predictive models and ensuring their alignment with criminological principles.

This thesis will delve deeper into the role of predictive modelling in law enforcement, exploring various methodologies and ethical considerations through real-world case studies. It aims to highlight how these models can support effective crime prevention while ensuring a commitment to justice and equity.

2.6.2 Predictive Modelling for Proactive Policing

The relevance of predictive modelling in today's policing landscape cannot be overstated. As crime patterns become more intricate and unpredictable, traditional reactive strategies often fail to meet the dynamic needs of modern communities. Law

enforcement agencies are increasingly turning to predictive analytics as a means of shifting from a reactive to a proactive approach, enabling them to anticipate criminal activity before it occurs and deploy resources with precision [115]. By leveraging historical crime data, statistical modelling, and machine learning algorithms, predictive policing provides law enforcement with an evidence-based framework for crime prevention, enhancing both efficiency and effectiveness in resource allocation. Empirical research has consistently highlighted the profound impact of predictive modelling on crime reduction. [116] demonstrated that data driven patrol strategies significantly reduce crime rates by focusing police presence in high risk areas identified through predictive analytics. Real world implementations further validate these findings, with law enforcement agencies across various jurisdictions integrating predictive models into their operational frameworks. In Los Angeles, the LAPD has successfully deployed predictive algorithms to identify crime hotspots, allowing for optimized officer deployment and improved community engagement [117]. By analyzing temporal and spatial crime patterns, this approach has not only enhanced the efficiency of patrol strategies but has also fostered stronger trust between police and local communities by demonstrating a commitment to proactive crime prevention. Similarly, the Chicago Police Department's "Predictive Policing Program" exemplifies the transformative potential of data driven law enforcement. Through the analysis of historical crime data, social networks, and behavioral patterns, the program anticipates potential incidents and identifies individuals at heightened risk of criminal involvement whether as offenders or victims [118]. This intelligence driven methodology facilitates timely interventions, enabling law enforcement to implement targeted strategies such as community outreach programs, rehabilitation initiatives, and pre emptive policing efforts. Beyond crime deterrence, these predictive approaches contribute to a broader shift in policing philosophy one that prioritizes prevention over reaction, strategic foresight over crisis response, and community well being over mere enforcement. Yet, the true significance of predictive modelling lies not just in its technical sophistication but in its capacity to redefine the fundamental principles of policing. By integrating big data analytics with human judgment, ethical considerations, and community collaboration, predictive policing offers a more balanced, transparent, and accountable approach to public safety. As advancements in artificial intelligence and data science continue to shape the future of law enforcement, predictive modelling stands as a

cornerstone in the evolution of modern policing one that has the potential to create safer, more resilient, and more just societies.

However, the implementation of predictive modelling is not without its challenges. While these technologies offer powerful tools for crime prevention, concerns surrounding algorithmic bias and the risk of reinforcing systemic inequalities have sparked significant ethical debates. Studies have shown that predictive algorithms, when trained on historical crime data, can inadvertently perpetuate biases embedded within past policing practices, disproportionately targeting marginalized communities [119]. For instance, an analysis of the predictive policing system used in Oakland, California, revealed that the algorithm disproportionately directed law enforcement to low-income, predominantly Black and Latino neighborhoods, despite similar crime rates in other areas that received less police scrutiny [120]. Such instances raise pressing concerns about the potential for over-policing and the erosion of public trust, particularly in communities that have historically experienced discriminatory law enforcement practices. Beyond geographic bias, predictive models that assess individual risksuch as Chicago's now-defunct Strategic Subject List (SSL) illustrate the dangers of algorithmic misclassification. Originally designed to identify individuals most at risk of involvement in gun violence, the SSL was criticized for its lack of transparency and its failure to account for socioeconomic factors, leading to individuals being labelled as high risk without clear justification [121]. In some cases, individuals flagged by the system reported experiencing increased police surveillance despite having no criminal record, further exacerbating tensions between law enforcement and the communities they serve. These examples underscore the ethical imperative for law enforcement agencies to ensure that predictive policing technologies do not reinforce existing disparities but instead contribute to fair and equitable policing practices. Therefore, while predictive modelling holds substantial promise for enhancing law enforcement efficacy, it must be implemented with careful oversight, transparency, and accountability. The development of bias mitigation techniques, such as algorithmic auditing, fairness aware machine learning, and community inclusive policy frameworks, is essential to preventing the unintended harms of predictive analytics [122]. Moreover, fostering public trust requires more than just technological refinement; it demands a commitment to ethical policing that prioritizes justice, inclusivity, and human rights. As law enforcement agencies continue to integrate data-driven strategies, the balance between innovation and ethical responsibility must remain

at the forefront to ensure that predictive policing serves all communities equitably, rather than reinforcing historical injustices.

2.6.3 Review of Studies on the Effectiveness of Predictive Modelling

Predictive modelling has emerged as a transformative tool in modern law enforcement, enabling agencies to shift from reactive policing to proactive crime prevention. Numerous studies have underscored its capacity to enhance policing outcomes, optimize resource allocation, and improve public safety [123]. By leveraging historical crime data, spatial analysis, and machine learning algorithms, predictive policing provides actionable intelligence that empowers law enforcement agencies to anticipate criminal activities and deploy interventions where they are most effective. The integration of predictive analytics into crime prevention strategies has not only led to notable reductions in crime rates but has also facilitated the efficient distribution of resources, ensuring that law enforcement efforts are both targeted and impactful. One of the most prominent applications of predictive modelling is hotspot prediction, where statistical algorithms analyze historical crime patterns to identify high-risk geographic areas. This approach enables police departments to allocate personnel and resources strategically, thereby maximizing crime deterrence. For instance, research by [124] demonstrated that predictive hotspot models significantly enhanced law enforcement's ability to direct patrols to areas with heightened criminal activity, leading to measurable crime reductions. Similarly, [125] found that predictive policing strategies in Los Angeles and Chicago resulted in improved police deployment and more effective crime suppression. By prioritizing these high-risk zones, law enforcement agencies can optimize their response efforts, reduce crime rates, and instill a greater sense of security within communities. Beyond geographical predictions, predictive modelling has been instrumental in offender profiling, allowing law enforcement agencies to identify individuals or groups more likely to engage in criminal behavior. These models incorporate a range of variables, including demographic characteristics, behavioral patterns, and environmental risk factors, to generate insights that support surveillance, intervention, and rehabilitation efforts. A study by [126] highlighted the efficacy of predictive analytics in offender profiling, demonstrating how data driven intelligence enhances case solving strategies and enables more precise targeting of high risk individuals. By refining offender profiling techniques, predictive modelling not only

aids in criminal investigations but also contributes to long-term crime prevention by addressing potential threats before they escalate.

In addition to its role in crime prediction and offender profiling, predictive modelling significantly improves resource allocation in law enforcement. By analyzing crime trends and risk factors, these models inform decision making processes regarding the strategic distribution of personnel, equipment, and funding. Research by [127] emphasized that predictive analytics driven resource allocation enhances police presence in crime prone areas, leading to more efficient law enforcement operations and improved public safety. This data-driven approach ensures that resources are utilized where they are most needed, preventing inefficient deployments and enhancing overall operational effectiveness. Despite its numerous advantages, predictive modelling is not without limitations. A key challenge lies in the reliability of the data used to train predictive algorithms. If the underlying crime data is incomplete, outdated, or inaccurately recorded, the resulting models may generate flawed predictions, leading to misdirected policing efforts. Furthermore, predictive models often rely on statistical correlations rather than causal relationships, meaning that while they can identify crime patterns, they do not necessarily explain the underlying social, economic, or environmental factors driving criminal activity [128]. Without a nuanced understanding of these factors, there is a risk of misinterpreting data and formulating ineffective policing strategies. Another critical concern is the potential overreliance on technology in law enforcement decision making. While predictive analytics provides valuable insights, it should serve as a complement rather than a replacement for human expertise. Law enforcement officers must critically interpret predictive outputs and integrate them with situational awareness and investigative experience to ensure that policing strategies remain adaptive and contextually appropriate. A purely data-driven approach, devoid of human oversight, may fail to account for the complexities of crime and societal dynamics, potentially leading to rigid or overly formulaic law enforcement practices. Nevertheless, when applied with diligence, transparency, and continuous refinement, predictive modelling offers immense potential to enhance crime prevention and policing efficiency. By integrating predictive analytics with informed decision-making, law enforcement agencies can significantly improve their ability to anticipate and prevent crime, allocate resources effectively, and foster safer communities. As advancements in artificial intelligence and data science continue to evolve, the refinement of predictive policing methodologies will play a pivotal role in shaping the

future of law enforcement, offering a smarter, more targeted, and more efficient approach to crime prevention.

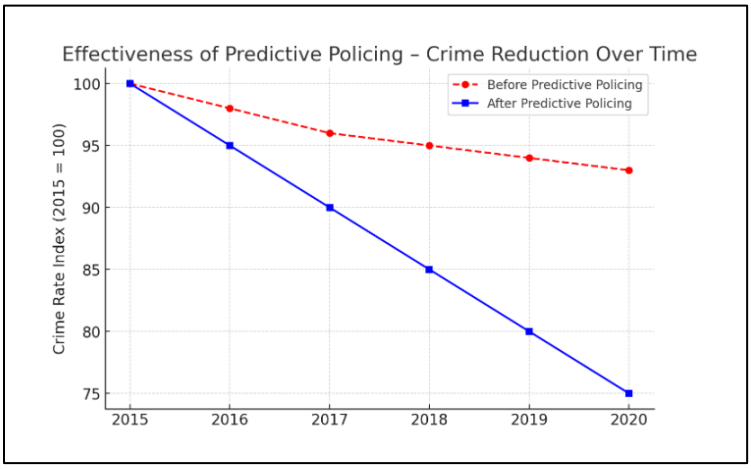


Figure 2.7 Crime Reduction Trends

2.7 Analysis of Case Studies and Real World Examples

Several compelling case studies illustrate the transformative power of predictive modelling in crime prevention, showcasing its potential to revolutionize law enforcement strategies through advanced analytics, machine learning algorithms, spatiotemporal modelling, social network analysis, and artificial intelligence (AI) driven forecasting**.** These methodologies enable more efficient resource allocation, targeted interventions, and proactive policing. The following examples highlight not only the successes achieved through predictive analytics but also the techniques employed and their broader implications for modern policing. One of the most widely recognized applications of predictive modelling is the Los Angeles Police Department's (LAPD) implementation of the PredPol system**,** which utilizes spatiotemporal analysis and machine learning to predict crime hotspots. This predictive policing program employs sophisticated algorithms that analyze historical crime data, focusing on the "three W's" of crime: where, when, and what type of crime is likely to occur [129]. The model is based on near-repeat theory**,** which suggests that crimes are likely to occur in spatial and temporal proximity to prior incidents**.** By applying Bayesian probability models**,** PredPol continuously refines its predictions, ensuring that law enforcement efforts are concentrated in high risk areas. According to Mohler et al., deploying officers proactively to these identified locations led to a measurable 20% reduction in property

crimes in certain districts during its initial implementation phase. Beyond crime reduction, the program also enabled the LAPD to optimize resource allocation through algorithmic decision-support systems, ensuring that officers were present in areas with the greatest potential for criminal activity. Additionally, the LAPD integrated community policing efforts alongside predictive modelling, using geospatial data visualization tools to engage local residents in discussions about safety, fostering greater trust and collaboration between law enforcement and the public.

Similarly, the Chicago Police Department (CPD) has pioneered the use of predictive modelling through its Strategic Subjects List (SSL), also known as the "heat list." The SSL employs machine learning classifiers, regression analysis, and social network analysis (SNA) to evaluate multiple risk factors, including prior arrests, known associates, geographic proximity to violent crime, and social ties within criminal networks. By applying predictive scoring algorithms, the system assigns risk scores to individuals most likely to be involved in violent incidents, either as victims or perpetrators [130]. Rather than focusing solely on enforcement, the CPD leveraged the predictive risk stratification model to engage with high risk individuals through targeted outreach programs, counseling services, and social interventions. The Urban Institute reported that neighborhoods where SSL was actively utilized saw reductions in gun violence and an increase in public perceptions of safety. The CPD also employed longitudinal impact assessments to evaluate the effectiveness of its interventions, ensuring that predictive policing was integrated with restorative justice frameworks rather than purely punitive approaches. The Seattle Police Department (SPD) provides another noteworthy example of predictive modelling in action, particularly in the context of gun violence prevention. The SPD employed real-time big data analytics, natural language processing (NLP), and sentiment analysis to track emerging threats. By analyzing real-time data sources such as crime reports, reported threats, and social media activity, SPD developed predictive models capable of identifying potential flashpoints for firearm-related crimes. The department used deep learning based threat detection algorithms to scan public data streams, identifying early warning signs of violent incidents. According to research conducted by the [131], the SPD's predictive approach resulted in a significant decrease in firearm related incidents, particularly in neighborhoods previously plagued by recurring gun violence. Additionally, reinforcement learning algorithms allowed SPD to optimize patrol strategies dynamically, ensuring that resources were deployed efficiently based on evolving crime

patterns. The integration of predictive analytics with real-time situational awareness tools, such as geofencing and automated alert systems, enhanced both operational efficiency and community safety.

In the United Kingdom, the Kent Police utilized predictive modelling through PredPol UK, an adaptation of the original PredPol algorithmic framework. The model applied random forest classification models and historical crime clustering techniques to analyze over 10 years of crime data, predicting potential burglary and theft incidents with high precision. According to [132], time series forecasting methods enabled the Kent Police to dynamically allocate patrols based on predicted burglary trends, resulting in a 15% reduction in reported burglaries in pilot areas. This initiative also demonstrated the potential of integrating predictive policing with community-based crime prevention programs, such as neighborhood watch groups, CCTV analytics, and predictive hotspot patrolling. The use of multi-layer perceptron (MLP) neural networks further enhanced predictive accuracy, refining crime forecasts through continuous learning from new crime incident reports. Another innovative example comes from the Netherlands, where the Dutch National Police employed artificial intelligence driven predictive modelling for urban crime prevention. Their system, known as the Crime Anticipation System (CAS), utilized AI based pattern recognition, weather-dependent crime prediction models, and event-based risk assessment algorithms to forecast crime risks in public spaces. CAS integrated structured and unstructured data sources, including socioeconomic indicators, public events, and surveillance data, to predict crimes such as pickpocketing and vandalism [133]. By deploying AI powered anomaly detection models, CAS was able to anticipate crime surges during large gatherings, concerts, and festivals, leading to a 13% decrease in reported petty theft incidents in monitored areas. The system's reliance on real-time data fusion techniques ensured that police could proactively adjust security measures, preventing criminal activity before it occurred.

These case studies exemplify the transformative impact of predictive modelling on law enforcement practices. While the successes underscore the efficacy of data driven approaches, they also highlight the necessity for ongoing evaluation and ethical considerations in the implementation of such technologies. Striking a balance between leveraging predictive analytics for public safety and safeguarding community trust remains a critical challenge for law enforcement agencies as they navigate the complexities of modern policing.

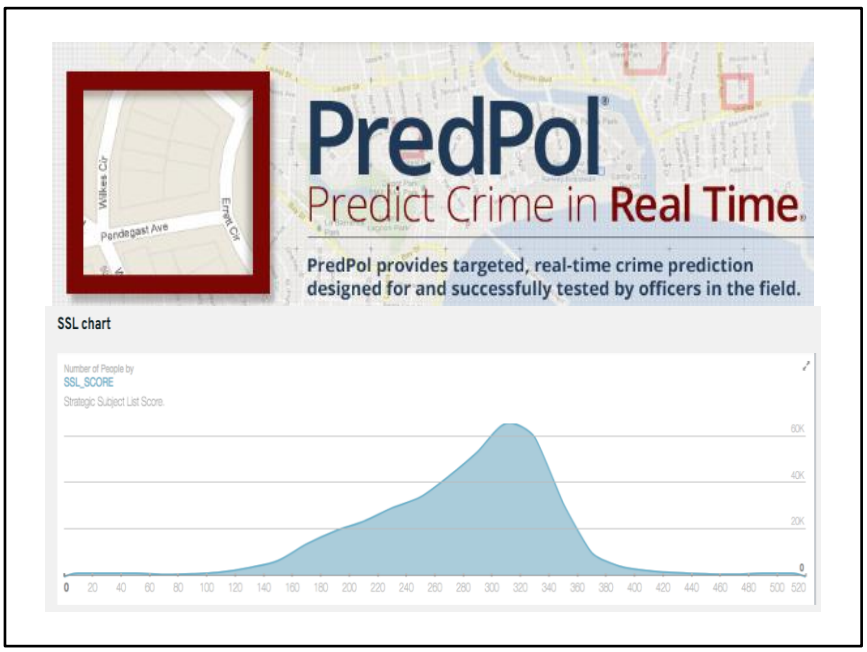


Figure 2.8 PredPol and SSL System used in the Real World

Table 2.3
Comparison of Predictive Policing Models and Their Impact on Crime Reduction

Police Department	Predictive Model Used	Methodologies	Predictive Model Used
Los Angeles Police Department (LAPD)	PredPol	Spatiotemporal analysis, Bayesian probability	20% reduction in property crimes
Chicago Police Department (CPD)	Strategic Subjects List (SSL)	Machine learning classifiers, social network analysis	Reduced gun violence, improved public safety
Seattle Police Department (SPD)	AI-driven gun violence prediction	NLP, sentiment analysis, reinforcement learning	Significant decrease in firearm-related incidents
Kent Police (UK)	PredPol UK	Random forest classification, time-series forecasting	15% reduction in burglaries
Dutch National Police	Crime Anticipation System (CAS)	AI-based pattern recognition, event-based risk assessment	13% decrease in petty theft incidents

2.8 Evaluation of Predictive Modelling Techniques

Predictive modelling has become a critical tool in modern law enforcement, enabling agencies to anticipate criminal activities and optimize resource allocation effectively. By leveraging data driven insights, law enforcement can move beyond reactive policing to proactive crime prevention strategies. Regression analysis is a foundational predictive modelling technique widely employed in criminology and law

enforcement. It enables researchers to examine relationships between multiple variables and crime occurrences, offering valuable insights into how different socioeconomic, demographic, and environmental factors influence crime rates. According to [134], regression models are particularly effective in identifying statistically significant predictors of crime, allowing law enforcement agencies to tailor intervention strategies accordingly. A key advantage of regression analysis is its interpretability unlike more complex models, regression outputs provide clear, quantifiable relationships between crime and influencing variables. This allows policymakers and law enforcement officials to make data driven decisions based on statistical significance rather than subjective assumptions. However, regression models are inherently limited by their dependence on historical crime data. They assume a linear relationship between variables, which may not always hold true in complex crime environments. Additionally, these models struggle to account for emerging crime trends, underreported crimes, or rapidly evolving criminal behaviors [135].

Machine learning algorithms have emerged as a powerful alternative to traditional statistical methods, offering greater flexibility and adaptability in crime prediction. Unlike regression analysis, machine learning can process vast amounts of structured and unstructured data, uncovering non linear relationships and previously undetected patterns. Techniques such as decision trees, random forests, support vector machines, and neural networks are widely applied in predictive policing. Research by [136] underscores the superiority of machine learning algorithms in crime forecasting due to their ability to continuously learn and improve as new data becomes available. These models are particularly effective in detecting hidden correlations and multifactorial interactions that may not be evident through traditional statistical methods. However, machine learning algorithms are often criticized for their “black-box” nature, which poses significant challenges in law enforcement applications [137]. Many models lack interpretability, making it difficult to justify policing decisions based on their outputs. This raises ethical concerns regarding transparency, accountability, and potential biases embedded in algorithmic decision-making. If trained on biased datasets, machine learning models can inadvertently reinforce systemic discrimination, leading to disproportionate targeting of certain demographic groups. Addressing these concerns requires rigorous model validation, bias mitigation strategies, and careful oversight to ensure ethical implementation in policing [138]. Spatial analysis, particularly through Geographic Information Systems (GIS), plays a critical role in

crime mapping and hotspot identification. By analyzing the spatiotemporal distribution of crime, GIS based models enable law enforcement agencies to deploy personnel strategically, optimizing patrol coverage in high-risk areas. As highlighted by [139], spatial analysis provides law enforcement with intuitive visual representations of crime trends, facilitating data driven decision making. One of the primary advantages of spatial analysis is its ability to incorporate environmental criminology principles. Techniques such as hot spot mapping, kernel density estimation, and risk terrain modeling (RTM) allow law enforcement to identify crime prone areas based on environmental attributes (e.g., poorly lit streets, abandoned buildings, or high foot traffic zones). These insights enable police departments to implement preventive measures such as increased surveillance, improved urban planning, or enhanced community policing initiatives [140]. Despite its effectiveness, spatial analysis has limitations. While it excels at identifying where crimes are likely to occur, it does not necessarily explain why they occur. Spatial models often fail to account for social, economic, and psychological factors influencing criminal behavior. Additionally, over reliance on geospatial predictions can lead to policing biases, where certain neighborhoods become over policed, exacerbating tensions between law enforcement and marginalized communities [141].

Advancements in big data analytics and real-time data processing are revolutionizing predictive policing, expanding the scope of crime forecasting beyond traditional datasets. Law enforcement agencies are increasingly integrating social media data, sensor based surveillance, and community reporting apps to enhance situational awareness and enable real time crime prediction. [142] highlight how these technologies allow police departments to respond proactively to unfolding incidents, improving emergency response times and community engagement. Moreover, predictive policing models are evolving to incorporate natural language processing (NLP) and sentiment analysis, enabling law enforcement to monitor public discourse for emerging threats, extremist rhetoric, or indicators of organized crime activity [143]. The Crime Anticipation System (CAS) developed by the Dutch National Police exemplifies this trend, using AI-driven anomaly detection to predict urban crime surges based on social and economic variables [144]. However, these advancements introduce new ethical and legal challenges. The use of real-time surveillance and social media monitoring raises concerns about privacy infringement, data security, and the potential for mass surveillance. Furthermore, real time predictive analytics require highly accurate data

streams inaccurate or biased data inputs can lead to misguided law enforcement actions, disproportionately affecting certain populations [145]. Predictive modelling has undeniable potential in transforming law enforcement by enabling proactive, data driven crime prevention strategies. Regression analysis provides interpretable insights into crime determinants but is constrained by its linear assumptions and reliance on historical data. Machine learning algorithms offer greater predictive power and adaptability but introduce challenges related to transparency and bias. Spatial analysis enhances crime mapping and resource allocation, yet it does not fully capture the social complexities underlying crime patterns. As predictive modelling evolves, the integration of big data, artificial intelligence, and real time analytics is shaping the future of law enforcement. While these advancements hold promise, they also necessitate rigorous ethical oversight to mitigate potential biases, ensure transparency, and uphold fundamental civil liberties. The responsible deployment of predictive analytics in policing requires a balanced approach, where technological innovation is complemented by ethical safeguards, community engagement, and evidence based policymaking. Predictive models offer significant potential to enhance law enforcement through proactive, data driven strategies. However, their effectiveness and public trust depend on responsible use. This requires ensuring data quality, reducing algorithmic bias, maintaining transparency, engaging communities, and upholding ethical and privacy standards. A balanced approach that integrates innovation with accountability is essential for sustainable and trustworthy predictive policing.

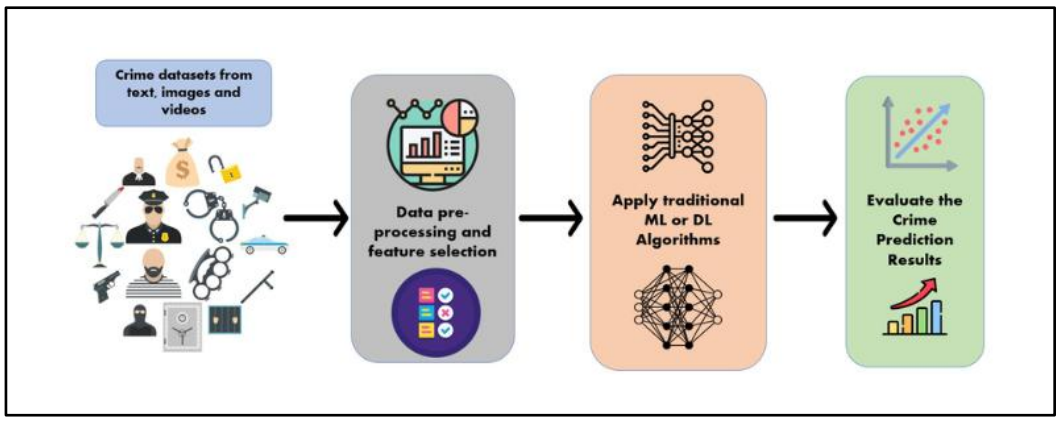


Figure 2.9 Workflow of Crime Prediction Using Machine Learning and Deep Learning Algorithms

Machine learning methods such as Decision Trees, Random Forest, and Neural Networks have become increasingly valuable in uncovering complex relationships

hidden within large and diverse crime datasets. Earlier studies, including that of [146], demonstrated that these algorithms can predict crime hotspots with considerable accuracy by linking historical crime data to social and economic indicators. Their findings showed that when such models were integrated into policing strategies, enforcement agencies could deploy resources more strategically, often leading to tangible improvements in crime prevention outcomes. Among the various approaches, the Random Forest algorithm has proven particularly effective for crime prediction tasks. As an ensemble method that aggregates multiple decision trees, it offers both high predictive accuracy and resilience against overfitting two qualities essential in dealing with noisy, real world data. Random Forest models are capable of capturing subtle nonlinear interactions between variables, allowing researchers to identify which contextual and socioeconomic factors such as unemployment rate, police visibility, or population density carry the greatest weight in shaping crime risk [147]. Beyond prediction, the model's ability to rank feature importance provides practical insights that can inform strategic planning and targeted interventions. Temporal modelling also plays a crucial role in understanding crime dynamics. [148] underscored that many types of crime exhibit cyclical or seasonal trends, where certain offences rise or fall predictably over time. By employing time series analysis, police departments can anticipate such fluctuations and adjust patrol schedules accordingly, improving their responsiveness to emerging risks. Spatial analysis further complements these predictive efforts by visualizing how crimes are distributed geographically. Studies such as [160] showed that combining Geographic Information Systems (GIS) with spatial modelling techniques helps reveal underlying spatial patterns, enabling police to focus their efforts on areas of persistent or emerging concern. Visualization tools like heat maps and point density plots make these patterns more accessible, allowing decision-makers to grasp high risk zones at a glance and allocate resources accordingly. Research consistently supports the effectiveness of integrating these predictive methods. For instance, [149] found that incorporating temporal dimensions such as time of day or day of week significantly improved prediction accuracy and informed tactical deployments during peak risk periods. Likewise, [150] and others have shown that ensemble based models like Random Forest outperform traditional statistical approaches by leveraging real time and multidimensional data sources. Importantly, the interpretability of Random Forest makes it well suited to law enforcement applications, as it not only predicts where and when crimes are likely to occur but also clarifies *why* certain areas are more vulnerable.

The operational impact of such predictive models has been examined in numerous studies. [151] reported that predictive policing systems frequently achieved 10–20% reductions in targeted crimes, while [152] confirmed that hotspot oriented policing informed by predictive analytics correlates with sustained decreases in local crime rates. Similarly, [153] highlighted that predictive frameworks, when properly implemented, allow agencies to optimize patrol routes and manpower distribution, improving both efficiency and public trust.

Nonetheless, these scholars also caution that the success of predictive models depends heavily on data quality, algorithm transparency, and ongoing calibration to prevent bias and ensure fairness. In essence, the integration of Random Forest and related machine learning models into spatial and temporal crime analysis represents a significant advancement in evidence-based policing. These tools extend beyond mere prediction they foster a deeper understanding of the structural and situational factors that sustain crime, enabling law enforcement to move from reactive response to proactive prevention. In addition to the approaches discussed above, recent studies have explored the application of advanced machine learning techniques such as Deep Learning for crime prediction. Deep learning models have demonstrated strong predictive performance in large-scale datasets with extensive training samples and complex feature representations. However, their effectiveness is often dependent on the availability of large volumes of structured data and extensive parameter optimization. Crime datasets collected in operational law enforcement environments frequently contain heterogeneous variables, missing values, and inconsistencies in reporting, which introduce noise and uncertainty into the analysis. Under such conditions, ensemble learning approaches such as Random Forest have been shown to provide stable and reliable predictive performance due to their robustness against noisy and incomplete data and their ability to model complex non-linear relationships. Furthermore, interpretability is an important requirement in law enforcement decision-support systems. Random Forest models provide feature importance measures that enable key contributing factors to be identified and interpreted by Investigation Officers. This transparency supports practical decision-making and strengthens confidence in system outputs. Based on these considerations and the characteristics of crime datasets typically encountered in policing environments, Random Forest represents a suitable predictive modelling approach for operational crime risk assessment systems such as ICRAS.

Table 2.4
Comparative Analysis of Selected Machine Learning Algorithms in Crime Prediction

Algorithm / Study	Core Principle	Strengths in Crime Prediction	Limitations / Challenges	Key Findings / Applications
Decision Tree (DT)	Uses conditional rules to split data hierarchically and classify outcomes	Easy to interpret; identifies key decision variables clearly.	Tends to overfit; less stable with complex or noisy datasets.	Suitable for identifying direct relationships between socio-economic factors and crime patterns.
Random Forest (RF)	Ensemble of multiple decision trees combined through bagging to improve accuracy.	High predictive accuracy; robust against overfitting; highlights feature importance.	Slightly reduced interpretability compared to a single tree; computationally heavier.	Achieved up to 90% accuracy in predicting crime hotspots; identified unemployment and police presence as key predictors.
Neural Network (NN)	Mimics human brain neurons to learn nonlinear relationships between variables.	Excellent at capturing complex patterns; adaptable to high-dimensional data.	Requires large datasets; lacks transparency ("black box").	Effective for detecting hidden spatial-temporal crime trends but less explainable to policymakers.
Support Vector Machine (SVM)	Finds the optimal boundary between classes by maximizing the margin in feature space.	Performs well on high-dimensional data; strong for binary classification.	Sensitive to parameter selection; less efficient for large-scale data.	Used to classify crime likelihood based on socio-demographic and spatial indicators with moderate success.

2.8.1 Selection of Predictive Model for ICRAS

Various machine learning techniques have been applied in crime prediction studies, including Support Vector Machines (SVM), Artificial Neural Networks (ANN), Decision Trees, and Deep Learning models. These techniques have demonstrated varying levels of predictive performance depending on dataset characteristics, feature complexity, and data quality. Deep learning approaches, in particular, have shown promising results in large-scale crime prediction studies where extensive training datasets are available. However, several studies have reported that deep learning models require substantial volumes of data and extensive parameter tuning to achieve stable performance. Crime datasets used in operational law enforcement environments often contain heterogeneous variables, missing values, and inconsistencies in reporting. Under such conditions, ensemble learning approaches such as Random Forest have been

shown to provide more stable and reliable predictive performance. Random Forest algorithms are capable of handling mixed data types and non-linear relationships while maintaining robustness against noise and incomplete data.

Previous studies have demonstrated that Random Forest often performs competitively or outperforms more complex models when applied to moderate-sized crime datasets. In addition to predictive performance, interpretability is an important requirement in law enforcement decision-support systems. Random Forest provides feature importance measures that enable the identification of key factors influencing crime risk predictions. This capability supports transparency and practical interpretation of results by Investigation Officers. Based on these considerations and the characteristics of the dataset used in this study, Random Forest was selected as the primary predictive modelling technique for the Integrated Crime Risk Assessment System (ICRAS).

2.9 Database Design in Crime Risk Assessment

A well-structured database serves as the core foundation of any effective crime risk assessment system, functioning as the central repository where diverse data sources are integrated, processed, and analyzed to produce actionable intelligence. In the contemporary law enforcement environment, where the volume and complexity of crime related information have expanded exponentially, the design of the database architecture directly determines the accuracy, reliability, and efficiency of crime risk analysis. An intelligently designed database is not merely a storage mechanism, it represents the conceptual and operational backbone that transforms fragmented data streams into coherent knowledge frameworks that support evidence based decision making. The success of any analytical model or predictive policing system is inherently dependent on how data is organized and processed within its database infrastructure. Crime data typically originates from multi sources data including criminal records, case reports, demographic data, surveillance systems, and public digital platforms. Each of these sources varies in format, granularity, and reliability. Without a structured and unified database design, these disparate datasets remain isolated, limiting the analytical capacity of law enforcement systems. Therefore, a relational and normalized database model is essential to ensure data consistency, reduce redundancy, and establish logical relationships among entities such as *suspects, incidents, locations, timeframes,* and *risk*

indicators. In the context of crime risk assessment, database design must not only support data storage but also facilitate real-time data processing and analysis. This is particularly critical when dealing with high-frequency data inputs from mobile devices, sensors, or automated reporting systems. As such, the architecture must incorporate data processing mechanisms as an integral component rather than a supplementary function. These mechanisms comprising extraction, transformation, validation, and storage enable the system to handle raw data dynamically, transforming it into structured and interpretable information ready for analytical modeling. The integration of data processing within the database architecture is what differentiates a static repository from an intelligent analytical system. By embedding automated data cleansing, normalization, and validation layers, the system ensures that only accurate and meaningful data enters the analytical pipeline. This minimizes errors, enhances the precision of risk prediction algorithms, and optimizes the use of computational resources. In essence, the database functions as both the heart and the filter of the crime risk assessment process, maintaining the balance between data accessibility, security, and analytical readiness.

Furthermore, the importance of data processing within the database extends to ethical and governance dimensions. As crime data often contains sensitive personal and demographic information, the database design must incorporate stringent access control, encryption, and audit mechanisms. These safeguards ensure compliance with data protection laws and uphold the ethical standards of modern policing, particularly in preventing misuse or bias in predictive analytics. From a strategic standpoint, a database that effectively integrates data processing capabilities empowers law enforcement agencies to transition from reactive crime reporting to proactive intelligence generation. Such a database supports complex querying, spatial temporal analysis, and the application of machine learning models that can uncover hidden correlations across datasets. More importantly, it facilitates inter-agency collaboration, enabling multiple departments or jurisdictions to share standardized, verified data for comprehensive risk assessments.

2.9.1 The Importance of Data Processing in Crime Risk Assessment and Its Framework

In the past, policing was primarily reactive law enforcement agencies responded to incidents only after they had occurred. However, with the rise of interconnected digital systems and the explosion of data from multiple sources, the need for a more proactive and predictive approach has become increasingly evident. Leveraging data for crime risk assessment allows agencies to move beyond traditional reactionary models and instead anticipate and prevent criminal activity before it happens. This transformation is not merely a matter of technological adoption; it represents a necessary evolution in policing practice, driven by the growing complexity of modern crime and the demand for more focused, data-informed interventions. Yet, this shift brings its own set of challenges. While the volume of available data has grown exponentially, the true difficulty lies not in collection but in managing, processing, and analyzing it effectively. The reliability of any crime risk assessment depends heavily on the quality of the underlying data. Today, law enforcement agencies draw information from an ever widening range of sources traditional police records, arrest logs, and case files, alongside newer forms such as social media content, surveillance feeds, and even sensor data generated by smart city infrastructures. This deluge of data offers significant potential for uncovering patterns and predicting trends, but it also introduces complexity, inconsistency, and a high risk of information overload if not properly processed. Raw data, in its initial form, is rarely fit for analytical use. It is often incomplete, duplicated, or inconsistent, and even minor errors can compromise the validity of predictive models. In crime risk assessment, inaccurate data can have particularly serious implications, leading to misguided policy decisions, misallocation of resources, or inequitable targeting of individuals or areas. Hence, the process of cleaning, validating, and standardizing data becomes a critical prerequisite to ensure that the analysis produces meaningful and ethically sound outcomes. Adding to this complexity, crime-related data rarely originates from a single source. Law enforcement databases may contain structured, verified records, while external sources such as social media or public reporting platforms often consist of unstructured or noisy information. Integrating these heterogeneous inputs into a coherent dataset suitable for analysis requires systematic data management and processing strategies. Without such mechanisms, valuable insights may remain buried in disorganized information, limiting

the potential for data-driven crime prevention. To address these issues, several structured frameworks have emerged to transform raw data into actionable intelligence. Among the most recognized are the Knowledge Discovery in Databases (KDD) process, the Cross-Industry Standard Process for Data Mining (CRISP-DM), Big Data frameworks such as Hadoop–Spark, and the Extract, Transform, and Load (ETL) architecture. Each framework represents a distinct approach to data processing. The KDD model emphasizes iterative refinement from data selection and preprocessing to transformation and evaluation enabling the identification of hidden patterns within large datasets [154]. While powerful for exploratory research, its cyclical and time-intensive nature limits its use in real-time operational settings. The CRISP-DM methodology, on the other hand, provides a structured, repeatable framework that begins with business understanding and continues through data preparation, modeling, and deployment [155]. Its cyclical design allows findings to inform and refine each stage, aligning technical outputs with organizational objectives. However, it still relies heavily on manual oversight and human interpretation.

In contrast, Big Data ecosystems such as Hadoop and Spark overcome many of these limitations by distributing data storage and computation across multiple systems, allowing for near real time processing of massive and complex datasets [156]. These frameworks support scalability and machine learning integration, though they require advanced technical infrastructure and data governance. Complementing these models, the Extract, Transform, and Load (ETL) framework offers a streamlined and operationally efficient solution tailored to large-scale, continuous data integration. It involves three sequential stages: extracting data from diverse sources; transforming it through cleaning, normalization, and validation; and loading it into centralized repositories for analysis. Within crime analytics, ETL acts as the backbone of modern data infrastructure integrating structured police databases with semi structured or unstructured external intelligence feeds [157]. By standardizing disparate datasets, the ETL process ensures analytical consistency and supports timely, evidence based decision making. Compared to KDD and CRISP-DM, which emphasize discovery and modeling for research purposes, Big Data and ETL frameworks prioritize automation, scalability, and interoperability qualities that are indispensable for operational law enforcement systems. In particular, ETL serves as a crucial link between data warehousing and predictive intelligence, providing a stable foundation upon which advanced crime risk models, such as the Integrated Crime Risk Assessment System

(ICRAS), can function effectively. Moreover, the ability to transform fragmented, inconsistent, and high volume data into coherent intelligence lies at the heart of modern crime risk assessment. Frameworks such as KDD, CRISP-DM, Hadoop–Spark, and ETL enable this transformation, but ETL stands out for its practicality, efficiency, and scalability [158]. Its emphasis on data integrity and process automation makes it a key enabler of proactive, data-driven, and ethically responsible policing in the digital age.

2.9.2 Principles of Database Design in Crime Analysis

Effective database design in the context of crime analysis encompasses several key principles that ensure the integrity, efficiency, flexibility, security, and usability of the database:

i. Data Integrity: Maintaining the accuracy, consistency, and reliability of data is fundamental to database integrity. Techniques such as data validation, normalization, and constraint enforcement are essential for preserving data integrity and preventing anomalies or errors.
ii. Efficient Storage: Crime databases often contain large volumes of data, necessitating efficient storage mechanisms to optimize performance and minimize storage costs. Techniques such as data compression, indexing, and partitioning help improve storage efficiency and retrieval speed.
iii. Flexibility: Crime databases must be designed to accommodate evolving data requirements and analytical needs. Flexible data models, schema designs, and data governance frameworks enable scalability and adaptability to change demands without compromising data integrity.
iv. Security: Protecting sensitive crime data from unauthorized access, modification, or disclosure is paramount. Role based access controls, encryption, and audit trails help safeguard data integrity and confidentiality, ensuring compliance with privacy regulations and ethical standards.
v. Usability: Crime databases should be user-friendly and intuitive, facilitating easy data retrieval, analysis, and reporting for law enforcement personnel and analysts. Intuitive user interfaces, query languages, and visualization tools

enhance usability and accessibility, enabling users to derive actionable insights from the data.

2.9.3 Database Design Principles

Existing literature provides valuable insights into database design principles specific to crime analysis, highlighting various considerations and best practices:

a) Data Model Selection: Researchers have explored different data models, such as relational, hierarchical, and object-oriented models, and their suitability for crime databases based on factors such as data complexity, query requirements, and scalability needs. For example, [159] introduced the Entity-Relationship (ER) model, providing a unified view of data that has since become widely adopted in database design.
b) Normalization Techniques: Several studies have investigated normalization techniques and their applicability to crime databases for reducing data redundancy and improving data consistency. [160] Pioneered normalization principles for relational databases, advocating for the elimination of data redundancies through normalization to enhance data integrity.
c) Indexing Strategies: Efficient indexing is essential for optimizing query performance and data retrieval speed in crime databases. Research has explored various indexing strategies, including B-tree, hash-based, and bitmap indexing. [161] provided comprehensive coverage of indexing techniques in database management systems, offering insights into their strengths and limitations in optimizing query performance.
d) Data Warehousing: Data warehousing techniques have gained prominence in crime analysis for consolidating and integrating data from disparate sources. Studies have examined data warehousing architectures, ETL processes, and OLAP tools for crime analysis applications. [162] discussed data warehousing concepts and methodologies, emphasizing the role of data warehouses in facilitating comprehensive and integrated analysis of crime data.
e) Spatial and Temporal Data Handling: Crime analysis often involves spatial and temporal data, requiring specialized database techniques for handling location-based and time-series data. Researchers have explored spatial indexing, spatial

query optimization, and temporal database models for crime analysis applications. [163] provided insights into spatial and temporal data mining techniques, addressing the challenges and opportunities in analysing spatial and temporal crime data for effective decision-making.

The way crime databases are built, organized, and managed plays a crucial role in helping law enforcement agencies make informed decisions and respond effectively to criminal activities. At the heart of effective crime analysis is the ability to integrate data from multiple sources whether its police reports, surveillance footage, or social media activity into a single, unified system. These centralized data hubs, often structured as data warehouses or lakes, provide a comprehensive view of crime patterns, making it easier for investigators to connect the dots and uncover meaningful insights.A well-organized database isn't just about storing large amounts of data it's about structuring that data in a way that makes it easy to analyze and use. Hierarchical databases, for example, allow law enforcement to break down crime statistics from a national level all the way down to specific neighborhoods. This kind of system makes it possible to track crime trends at different levels starting with overall national crime rates, then zooming in to see which states have the highest levels of certain offenses, then drilling down to citywide, district, and even street-level data. A police department using such a system could, for instance, identify that burglaries are on the rise in a particular neighborhood and allocate more patrols accordingly. Beyond simple hierarchies, graph-based database structures take things a step further by mapping relationships between people, locations, and criminal activities. This is especially useful for tracking gang affiliations, fraud networks, or repeat offenders who operate across different jurisdictions. These systems allow law enforcement to identify key players in criminal organizations, predict where crimes might happen next, and take proactive measures to prevent them.Of course, a crime database is only as good as the accuracy of the information it contains. If the data is full of errors or missing key details, it can lead to bad decisions. That's why quality control is so important. Automated systems that clean, validate, and verify data help ensure that crime reports, forensic records, and suspect profiles are reliable and up to date. With accurate data, law enforcement agencies can make confident decisions based on facts, not guesswork.

As crime data continues to grow, scalability is becoming more important than ever. Today's crime databases need to be able to handle massive amounts of information without slowing down. Whether through cloud based systems or distributed databases that can expand as needed, agencies must ensure that their systems remain fast, efficient, and capable of delivering real-time insights. A database that can scale effectively means officers and analysts can access critical crime data whenever they need it, without delays or technical bottlenecks. Another key component of crime databases is metadata essentially, data about the data. Metadata provides important context, such as timestamps, locations, crime categories, and case statuses, helping investigators make sense of the vast amounts of information at their disposal. For example, metadata can help identify crime trends by linking incidents to specific times of day, days of the week, or even economic conditions. A well-maintained metadata system can also enhance coordination between different agencies, ensuring that information flows seamlessly between police departments, forensic labs, and the judicial system. By managing metadata effectively, law enforcement can transform raw data into clear, actionable insights that improve crime prevention and public safety.

2.9.4 Data Sensitivity and Secrecy in Predictive Policing

The integration of predictive models and crime databases into law enforcement operations has introduced significant improvements in crime prevention. However, it also involves the handling of highly sensitive and classified information. In Malaysia, crime-related data such as criminal records, investigation reports, surveillance outputs, and strategic intelligence is often protected under the Akta Rahsia Rasmi 1972 (Official Secrets Act). This act prohibits the unauthorized disclosure of official information and enforces strict protocols for data access and usage. Researchers who require access to these classified datasets must undergo formal vetting and sign confidentiality agreements, acknowledging their responsibility to maintain secrecy and prevent any form of data leakage. This approach is necessary to protect ongoing investigations, operational strategies, and national security interests. It also ensures that the integrity and confidentiality of police work remain uncompromised. Unlike open source datasets used in general academic research, law enforcement data cannot be freely accessed or published. The use of predictive technologies in this context must therefore be balanced with robust data protection measures. Even anonymized data can pose risks if patterns

or sensitive insights are inadvertently exposed. As such, any form of data sharing must be strictly controlled and governed by official protocols. The development and application of predictive policing tools must also consider the need for internal data protection standards, access restrictions, and non-disclosure policies. Effective collaboration between law enforcement agencies and researchers should be guided by clearly defined boundaries that prioritize data security over transparency, ensuring that innovation does not come at the expense of confidentiality or public safety.

2.9.5 Human Factors in Crime Prevention

In addition to ethical considerations, the use of technology in law enforcement is subject to a complex web of legal frameworks and regulations in Malaysia. These legal frameworks govern various aspects of technology implementation, including data collection, surveillance, and the use of predictive models. For example, the Federal Constitution of Malaysia guarantees fundamental liberties, including the right to privacy, which influences the regulation of surveillance activities by law enforcement agencies. Furthermore, privacy laws and regulations play a crucial role in governing the collection, storage, and use of data in crime databases and predictive models. The Personal Data Protection Act 2010 (PDPA) sets out principles for the processing of personal data in commercial transactions, which also impact law enforcement's handling of personal information. Additionally, laws such as the Communications and Multimedia Act 1998 (CMA) regulate electronic surveillance and cyber-related crimes, ensuring lawful and ethical use of digital tools in crime prevention. Moreover, legal frameworks governing the use of predictive models in law enforcement aim to ensure fairness and accountability in decision-making processes. The Malaysian Penal Code and the Evidence Act 1950 outline provisions to prevent discrimination and ensure that data-driven decisions adhere to principles of justice and fairness. The Independent Police Conduct Commission Act 2020 (IPCC) was also established to oversee and investigate police misconduct, promoting transparency in law enforcement operations. We can conclude that the integration of technology in law enforcement presents both opportunities and challenges, particularly concerning crime databases and predictive models. Ethical considerations surrounding privacy, fairness, and accountability must be carefully addressed to ensure that technology is used responsibly and ethically in law enforcement practices. Moreover, compliance with Malaysian legal frameworks and

privacy regulations is essential to protect individuals' rights and liberties while leveraging technology to enhance public safety. Moving forward, policymakers, law enforcement agencies, and technology developers must collaborate to develop ethical and legal frameworks that uphold democratic values and safeguard individual privacy rights in the digital age.

2.10 Similar Implementation of Technology

The integration of technology in law enforcement has significantly transformed the landscape of crime prevention and investigation. Several success stories illustrate how various jurisdictions have effectively harnessed technological innovations to enhance public safety. For instance, the Los Angeles Police Department's (LAPD) use of the predictive policing tool "PredPol" has led to a notable reduction in property crimes by as much as 30% in certain areas [164]. By leveraging algorithms that analyze crime patterns, officers can strategically allocate resources where they're most needed, demonstrating how technology can play a proactive role in addressing criminal activity. Moreover, the establishment of real-time crime centers (RTCCs) has redefined law enforcement operations. The Chicago Police Department’s RTCC, which integrates data from various sources such as surveillance cameras and gunshot detection systems, provides officers with immediate insights during critical incidents [165]. This enhanced situational awareness allows for quicker, more informed responses to emerging threats, showcasing the potential of technology to improve law enforcement efficacy. However, the journey toward technological integration is not without its challenges. One significant concern is the risk of reinforcing biases and discrimination. Predictive policing models, while useful, often rely on historical data that can reflect systemic inequalities, raising ethical questions about their application [166]. Additionally, the introduction of body worn cameras intended to boost transparency and accountability faces challenges related to privacy and the management of recorded footage [167] To effectively address these issues, law enforcement agencies must establish clear policies and ethical guidelines for technology use. Engaging with community members and emphasizing transparency can help alleviate concerns while building trust between law enforcement and the communities they serve. By examining these success stories and challenges, it becomes evident that a balanced approach to technological integration is

crucial. This ensures that the benefits of innovation are realized without compromising ethical standards, ultimately contributing to safer communities.

The table clearly illustrates how the strategic integration of advanced technology can revolutionize law enforcement practices. Malaysia has a unique opportunity to leverage these innovations to enhance crime detection, optimize resource allocation, and strengthen community relations through increased transparency and accountability. By embracing data-driven and real time monitoring approaches, Malaysian law enforcement can significantly improve public safety outcomes, foster greater trust within communities, and establish a proactive framework that addresses crime more effectively and efficiently**.**

Table 2.5
Examples of Technological Innovations in Modern Law Enforcement

Case Study	Description	Impact
Predictive Policing in Los Angeles	The LAPD implemented predictive policing algorithms to identify crime hotspots and allocate resources strategically. Advanced data analytics and machine learning techniques were used.	Led to a notable reduction in crime rates, demonstrating the effectiveness of technology in augmenting traditional law enforcement practices.
Body worn Cameras in New York City	The NYPD introduced body-worn cameras to enhance transparency and accountability in law enforcement. These cameras provided firsthand evidence of police interactions.	Reduced complaints against officers, improved community relations, and fostered trust and accountability in policing.
Real-time Crime Centres in Chicago	The CPD established real-time crime centres equipped with surveillance technology and data analytics to monitor crime in real-time, analyze trends, and coordinate response efforts.	Improved crime detection and prevention, resulting in safer communities and enhanced public safety.

2.11 Critical Gap Analysis

This Critical Gap Analysis compares existing crime risk assessment systems with the proposed Integrated Crime Risk Assessment System (ICRAS) to highlight the need for a more robust, data-driven, and scalable approach to crime risk management. The analysis identifies key limitations in current systems and demonstrates how ICRAS addresses these gaps, offering a more effective solution. Crime risk assessment systems play a crucial role in law enforcement by enabling proactive policing, improving resource allocation, and predicting crime trends. However, existing systems face

significant limitations. Many rely on fragmented data and standalone tools, which hinder their ability to provide comprehensive assessments. These systems often use historical crime data without integrating other critical factors, such as socio-economic or environmental variables, limiting their predictive capabilities. Additionally, current systems experience delays in decision-making, reducing their effectiveness in addressing emerging crime trends. Furthermore, predictive models in existing systems are often based on basic statistical methods, which lack the adaptability and precision of modern machine learning techniques.

The proposed ICRAS overcomes these limitations by integrating multiple data sources into a centralized platform, incorporating real-time data from crime records, socio-economic indicators, and environmental factors. This integration provides a more holistic and accurate risk assessment. ICRAS also employs the Random Forest algorithm for advanced predictive analytics, offering real-time crime predictions and proactive policing capabilities. Unlike current systems with limited user interfaces, ICRAS provides an intuitive, interactive dashboard, improving usability for officers. Moreover, ICRAS ensures data privacy and security by incorporating encryption, access controls, and compliance with privacy laws, addressing concerns seen in existing systems. The following table provides a summary of the critical gaps identified in the literature, comparing them to the capabilities offered by ICRAS. This table justifies the need for the development of a comprehensive, integrated crime risk assessment framework.

Table 2.6
Critical Gap Analysis of Existing Crime Risk Assessment Systems vs. Proposed ICRAS

Aspect	Existing systems / Apps	Proposed ICRAS	Gap Identified
Data Integration	PredPol (Uses historical crime data, lacks multi-source integration) COMPAS (Relies on actuarial data, limited integration of external data)	Centralized platform integrating multi source data (crime records, socio economic indicators, environmental data)	Lack of real time, centralized data integration in existing systems.
Predictive Capability	PredPol (Uses historical crime data to predict hotspots, limited predictive scope)	Predictive analytics using Random Forest algorithm for real time crime prediction and proactive policing.	Existing systems lack advanced predictive models for proactive policing.

Aspect	Existing systems / Apps	Proposed ICRAS	Gap Identified
User Interface	COMPAS (Basic interface for risk scoring, not user-friendly) PredPol (Maps crime hotspots but lacks interactivity)	Intuitive, interactive dashboard with real time visualizations and actionable insights.	Existing systems offer limited or non-interactive user interfaces, while ICRAS provides a user-centric interface that enhances interaction and decision-making.
Accuracy of Crime Risk Assessments	PredPol (High error rate in predictions, particularly in socio-economically disadvantaged areas) COMPAS (Criticized for racial bias and lack of transparency)	Data driven assessments using quantitative models (Risk Appetite Index, Crime Risk Matrix) and qualitative factors.	Higher accuracy and reduced bias in risk assessments with ICRAS using data-driven models that factor in socio-economic and environmental variables.
Technological Adaptability	PredPol (Uses basic statistical methods, lacks machine learning integration)	Utilizes machine learning, data analytics, and Random Forest algorithm for enhanced analysis and predictions.	Lack of machine learning integration in existing systems limits their ability to scale and adapt to new crime patterns, whereas ICRAS incorporates advanced machine learning for dynamic risk assessments.
Data Privacy and Security	COMPAS (Concerns over data privacy and security, particularly with sensitive offender data) PredPol (Limited security features for sensitive crime data)	Enhanced security with data encryption**,** access control, and compliance with data privacy laws.	Existing systems lack robust security features; ICRAS emphasizes enhanced security and compliance with privacy laws to safeguard sensitive data.

CHAPTER 3
RESEARCH METHODOLOGY

3.1 Research Design

The research design serves as the methodological foundation for achieving the objectives of this study through the development and evaluation of the Integrated Crime Risk Assessment System (ICRAS). A structured research design is essential to ensure that the selected methods align with the research objectives and support reliable analysis and interpretation of crime risk patterns [168]. This research introduces a novel, data-driven approach to crime risk assessment through the development of ICRAS, which integrates real-time data from multiple sources, such as historical crime data, socio-economic factors, and geospatial data. This novel integration enables dynamic crime predictions, which sets ICRAS apart from traditional systems that rely solely on static crime data. By shifting from reactive policing to proactive crime prevention, ICRAS offers timely, actionable insights for law enforcement agencies.

In the context of crime risk assessment, an effective research design must combine analytical modeling with practical system development to address the complexity of crime-related data and operational requirements [169]. This study adopts a mixed-methods approach that integrates quantitative analysis with qualitative evaluation to support the development of the proposed ICRAS framework. The mixed-methods design enables the study to combine empirical modeling of crime risk with practical assessment of system usability and operational relevance. The quantitative component focuses on analyzing crime patterns and developing predictive models using structured datasets. Statistical analysis and machine learning techniques (e.g., Random Forest algorithms) were applied to identify relationships between crime incidents and selected predictor variables, providing empirical support for the development of the Risk Appetite Index, Crime Risk Matrix, and predictive modeling components. Quantitative approaches are essential for producing reliable and evidence-based findings that support crime risk evaluation [170].

The qualitative component complements the quantitative analysis by incorporating feedback from Investigation Officers (IOs) and relevant stakeholders to ensure that the developed system meets operational requirements. User feedback and

expert input were used to refine the system design and evaluate its usability in practical law enforcement environments. This approach ensures that the analytical framework is not only technically valid but also suitable for real-world application [171].

The research design supports the implementation of the proposed ICRAS approach, which integrates data processing, risk scoring, and predictive modeling into a unified framework. The approach consists of four main stages: data integration using an Extract, Transform, and Load (ETL) process; structured risk evaluation using the Risk Appetite Index and Crime Risk Matrix; predictive modeling using the Random Forest algorithm; and visualization through a user-accessible platform. These components collectively support systematic crime risk assessment and enable Investigation Officers to interpret crime risk information more effectively. Ethical considerations were incorporated throughout the research process to ensure responsible data usage and participant protection. Ethical guidelines were followed to maintain confidentiality and informed consent, particularly in the handling of operational crime data and stakeholder feedback. The study complies with relevant Malaysian regulations, including the Personal Data Protection Act 2010, as well as internationally recognized research ethics standards. These measures ensure that the research was conducted in a responsible and ethically sound manner.

Stakeholder engagement formed an important part of the research design, particularly in the development and evaluation of the ICRAS platform. Collaboration with law enforcement personnel ensured that the system was aligned with operational needs and practical constraints. Participatory approaches have been shown to improve the applicability of research outcomes and support effective implementation in real-world environments [172].

This research design therefore provides a structured approach for developing and evaluating the ICRAS framework by integrating quantitative modeling, qualitative evaluation, and system implementation. The design supports the achievement of the research objectives by ensuring that the proposed approach is both analytically rigorous and operationally relevant. The novelty of ICRAS lies in its integration of real-time data, machine learning algorithms, and predictive analytics, which contribute to more proactive crime risk management, offering law enforcement agencies the tools needed to anticipate and mitigate crime in a timely and efficient manner.

3.1.1 Mixed Methods Research Design

Mixed methods research involves the systematic integration of quantitative and qualitative data collection and analysis within a single study or a coordinated set of studies. This approach is particularly effective for addressing complex research questions where single-method approaches may fall short. It allows for the convergence of different types of data to provide a richer, more detailed understanding of the research problem [173]. The quantitative component typically involves the collection and statistical analysis of numerical data. This phase is crucial for identifying patterns, trends, and relationships within large datasets. In the context of ICRAS, this involves analysing crime data to identify the action need to be taken to address that specific crime. The qualitative component, on the other hand, involves the collection and thematic analysis of non-numerical data, such as interview transcripts, focus group discussions, and observational notes. This phase provides deeper insights into the contextual and experiential aspects of the research problem, such as the lived experiences of victims, the perspectives of law enforcement officers, and the community's perception of crime and safety. The rationale for employing a mixed methods design in this study begins from the complex nature of crime and its multifaceted impacts on society. Crime is not only a statistical phenomenon but also a social issue that affects individuals and communities in profound ways. Therefore, understanding crime patterns and designing effective prevention strategies requires both quantitative and qualitative insights.

i. Quantitative Insights: Quantitative data allows for the identification of broad patterns and trends in crime. By analysing large datasets, it is possible to determine which areas are most affected by crime, what types of crimes are most prevalent, and how crime rates change over time. This information is crucial for resource allocation, policymaking, and the development of targeted interventions. [174].
ii. Qualitative Insights: Qualitative data, on the other hand, provides a deeper understanding of the contextual factors that influence crime and its impact on individuals and communities. Through interviews and focus groups, we can gain insights into the experiences of victims, the challenges faced by law enforcement officers, and the community's perception of safety. These insights

are essential for designing interventions that are not only effective but also sensitive to the needs and concerns of the community [175].

In developing the Integrated Crime Risk Assessment System (ICRAS) for the Royal Malaysian Police, combining quantitative and qualitative insights creates a well-rounded understanding of crime patterns and their impact. Quantitative data provides the foundation for identifying broad patterns and trends in crime. By analysing historical crime data from police records, we can uncover key information such as which areas are most affected, what types of crimes occur most frequently, and how crime rates shift over time. For instance, analysis might reveal that burglaries are more prevalent in specific neighbourhoods or that incidents of vehicle theft often increase during weekends. Additionally, statistical models can help link crime rates to underlying socioeconomic factors, such as unemployment levels or population density, giving decision-makers a clear view of the risks at play. These insights directly inform strategies like focusing law enforcement resources on high risk areas or tailoring policies to address specific vulnerabilities. At the same time, qualitative data provides a much needed human perspective, offering a deeper understanding of the factors driving crime and its effects on people and communities. Interviews and focus group discussions with police officers, community leaders, and residents reveal stories and experiences that numbers alone cannot capture. For example, law enforcement officers might share the challenges they face, such as limited manpower or logistical issues that affect response times. Victims of crimes like burglary or assault can describe the emotional and financial stress they endure, highlighting gaps in victim support systems. Similarly, feedback from residents might shed light on concerns about neighbourhood safety, like the need for improved street lighting or more visible police patrols. These perspectives help shape interventions that are not only effective but also sensitive to the lived realities of the people they aim to protect.

By bringing these two (2) approaches together, this research ensures that ICRAS is both data-driven and empathetic. Quantitative analysis lays the groundwork by identifying trends and risks, while qualitative insights provide the context needed to interpret these findings and make them actionable. This balanced approach doesn't just create a system that assesses crime risk accurately it builds one that resonates with the needs and priorities of the people it serves. In doing so, ICRAS becomes a tool that not

only guides law enforcement strategies but also fosters trust and collaboration between the police and the community, paving the way for safer, more resilient neighbourhoods

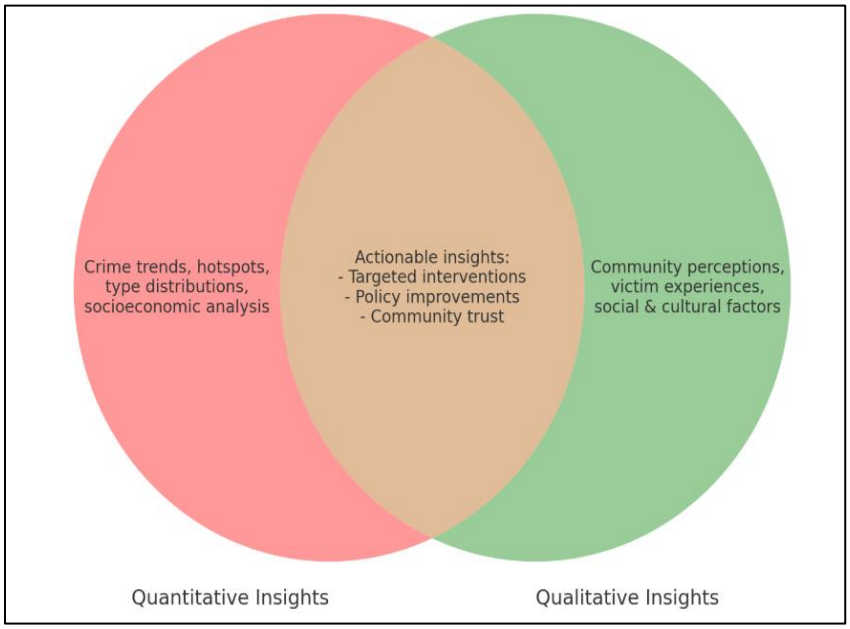


Figure 3.1 Integration of Quantitative and Qualitative Insights for Enhanced Crime Risk Assessment and Decision Making

3.1.2 Research Design of ICRAS

The research design for the Integrated Crime Risk Assessment System (ICRAS) adopts a mixed method approach, involving both qualitative and quantitative methodologies to ensure a comprehensive analysis and development process. This multifaceted approach is meticulously structured to integrate theoretical frameworks, empirical data, and technological advancements to create a robust, user cantered system for crime risk assessment and prevention. The research design is delineated into several stages.

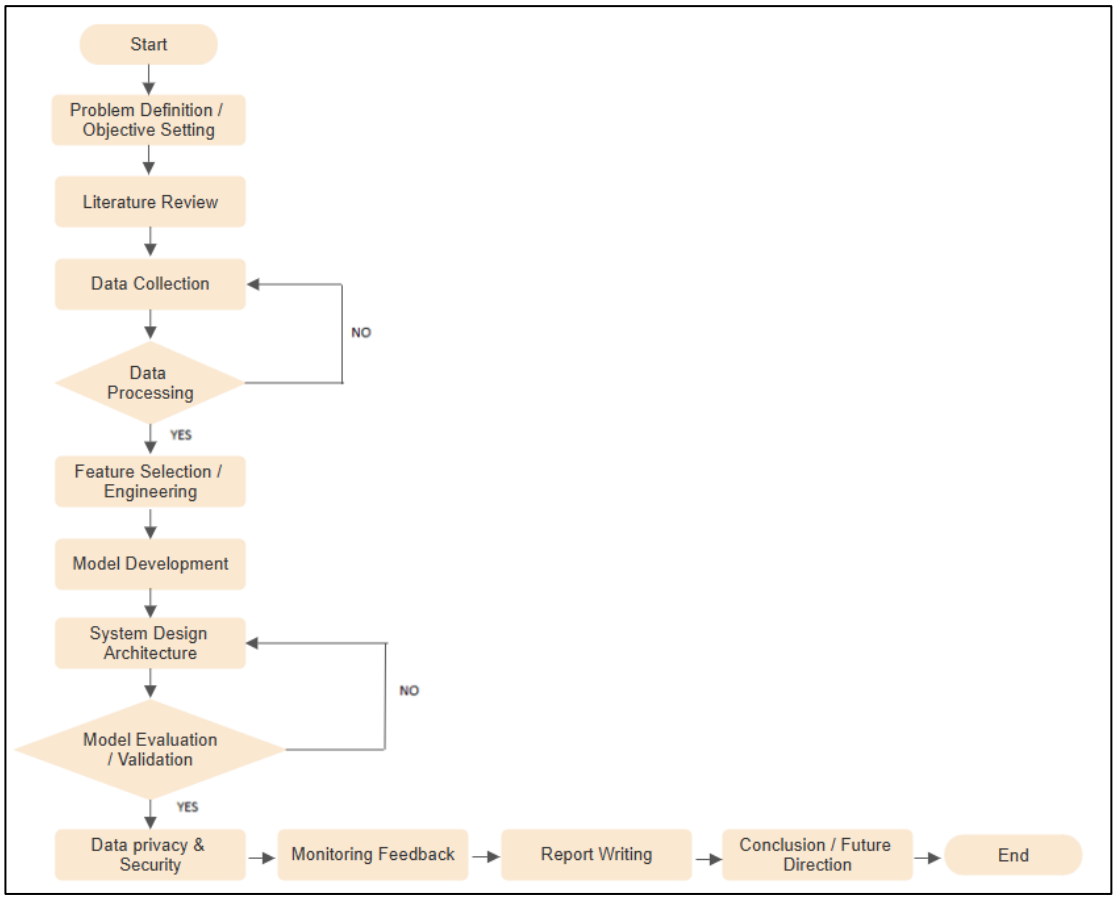


Figure 3.2 ICRAS Research Design Framework

The research design for the Integrated Crime Risk Assessment System (ICRAS) begins with problem definition and objective setting, where the primary aim is to address the absence of a comprehensive crime risk assessment framework in Malaysia. This initial phase involves recognizing the need for a systematic approach to predicting and mitigating crime risk, and establishing clear objectives for developing an advanced, data-driven system. The goal is to create a model that can provide actionable insights for law enforcement to enhance crime prevention and intervention strategies. This foundational phase is critical, as it defines the problem at hand and sets the direction for the entire research process. Following the problem definition, the literature review phase thoroughly examines existing studies on predictive policing, crime risk models, and data analytics in law enforcement. Given the lack of an established crime risk assessment system in Malaysia, this stage emphasizes the importance of evaluating international models and methodologies that have been successfully applied elsewhere. By identifying key gaps and opportunities, the research ensures the proposed system is both innovative and tailored to Malaysia's unique socio-economic and geographical landscape. Subsequently, data collection involves gathering comprehensive datasets

crucial to building the ICRAS model. This process integrates diverse data sources, including crime statistics, socio-economic indicators, demographic information, and environmental factors. Since Malaysia currently lacks a crime risk assessment framework, acquiring accurate and representative data is vital to constructing a reliable predictive model that reflects the complexity of local crime dynamics. The raw data collected undergoes data processing, where it is cleaned, organized, and transformed into an analysis-ready format. This includes addressing missing values, resolving inconsistencies, and standardizing variables. Following data processing, feature selection and engineering identify and create the most relevant and impactful variables, enhancing the predictive power of the model by focusing on factors most influential to crime risk within the Malaysian context. The model development phase serves as the core of ICRAS, exploring various machine learning algorithms and predictive techniques to determine the most effective method for assessing crime risk. Techniques such as decision trees, random forests, and neural networks are tested to identify the best approach for accurately predicting crime hotspots and individual risk factors. Given the pioneering nature of this system in Malaysia, this stage is pivotal to developing a tailored and robust solution that meets the specific challenges faced by Malaysian law enforcement. Following model development, the system design architecture phase translates the predictive model into a fully functional crime risk assessment platform. The architecture is designed for seamless integration with existing databases, real time data processing, and user-friendly interfaces that provide actionable insights. Scalability and adaptability are prioritized to ensure the system evolves alongside new crime data and changing operational requirements, making it practical for sustained use in real world law enforcement settings. The subsequent model evaluation and validation phase rigorously tests the system's predictive accuracy and reliability using historical crime data and standard performance metrics, including accuracy, precision, recall, and F1-score. Cross validation ensures the model's generalizability to new data, guaranteeing robustness and dependable performance under varied conditions. Given the sensitive nature of crime related data, the next critical phase canters on Data Privacy and Security. This stage ensures that all data handling complies with stringent privacy standards and legal regulations to protect individuals' rights. The system incorporates advanced security protocols to safeguard data from unauthorized access and breaches, while transparent data governance practices promote accountability. Special attention is given to address sensitive information and implementing measures that prevent misuse or

discrimination stemming from algorithmic outputs. By prioritizing data privacy and security, ICRAS not only protects stakeholders but also builds trust among users and the public, which is essential for the system's ethical and sustainable deployment. Once deployed, the monitoring and feedback phase continuously tracks the system's performance, gathering input from law enforcement officers and other end-users to refine and adapt ICRAS. This iterative process ensures the system remains effective and responsive to emerging crime trends and operational challenges.

The research culminates in the report writing phase, where methodologies, system design, validation outcomes, and data privacy measures are comprehensively documented. This report also addresses limitations, practical implications for Malaysian law enforcement, and recommendations for further research and system enhancement. Finally, the conclusion and future direction phase reflects on the successes and challenges of the ICRAS project, proposing potential expansions such as broader geographical coverage, incorporation of additional data sources, and adoption of advanced predictive techniques. This ensures ICRAS continues to evolve as a cutting-edge tool for improving crime risk assessment and prevention in Malaysia.

3.1.3 Research Phase

The development of the Integrated Crime Risk Assessment System (ICRAS) follows a structured research process that combines both quantitative and qualitative methods. This mixed-method approach provides a balanced framework for understanding the complex dynamics of crime while ensuring that the system remains data-driven and contextually relevant. Quantitative methods emphasize structured data analysis, predictive modelling, and measurable performance indicators to identify patterns and guide evidence-based decision making. Complementing this, qualitative methods capture stakeholder perspectives, operational challenges, and ethical considerations, helping to refine the system's design and usability. Together, these methods form a cohesive framework that integrates empirical accuracy with practical insights. The research process is carried out through several key phases data system architecture, system design and development, testing and compliance, deployment, and sustainability each contributing to the creation of a robust and adaptable crime risk assessment model.

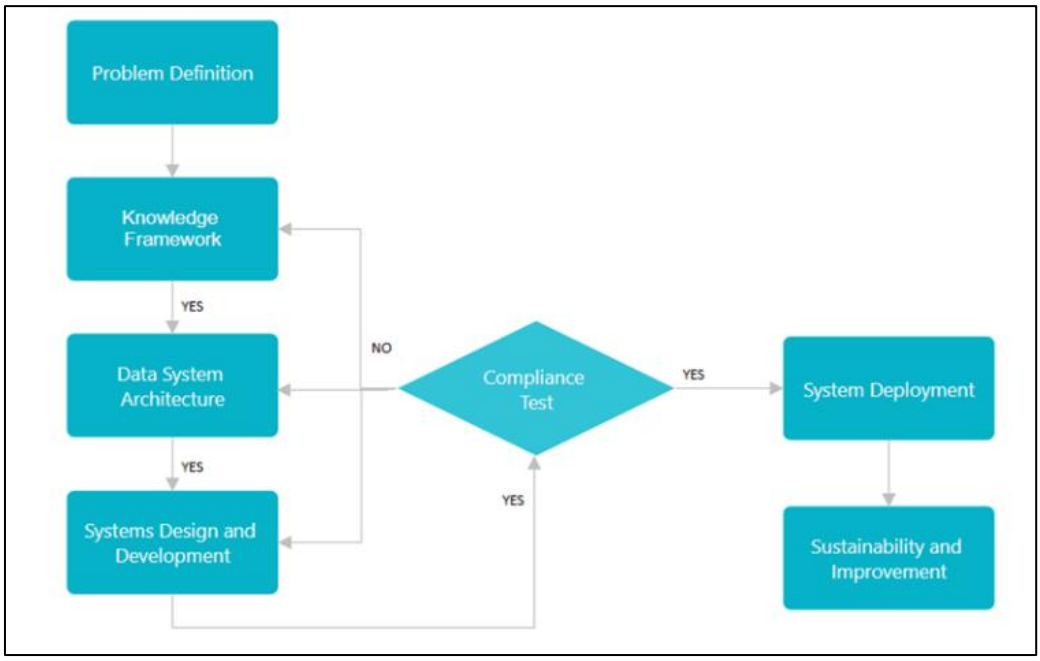


Figure 3.3 Process Developments of ICRAS

The initial stage, Problem Definition**,** involves a thorough examination of the challenges in crime risk assessment within the Malaysian context. It aims to define clear research objectives that address these challenges by enhancing decision-making processes, optimizing resource allocation, and establishing robust crime prevention strategies. This phase also includes a detailed review of existing national and regional crime prevention policies to ensure the system aligns with legal and ethical standards. Following this, the Knowledge Framework phase is crucial in establishing the knowledge framework that underpins the research. This includes an in depth literature review, identifying gaps in the current body of work and contextualizing the research within a global, regional, and local perspective. The goal is to ensure that the system is not only technically sound but also culturally sensitive and aligned with the socio-legal environment of Malaysia. By integrating both theoretical concepts and practical insights, this stage lays the groundwork for the subsequent phases. In the Data System Architecture phase, the focus shifts to the design and structuring of the system's data architecture. This involves the creation of an integrated, scalable database that consolidates crime data, socio-economic factors, and other relevant datasets. The data collection process involves both qualitative and quantitative methods, ensuring that the architecture supports advanced analytical processes such as predictive modelling. The system is designed to handle diverse data sources while maintaining integrity and accuracy, which are crucial for generating reliable crime risk assessments. The next

phase, System Design and Development**,** focuses on building the core functionalities of the system. This stage is where the system's front-end and back-end components are developed, ensuring seamless data processing, analysis, and visualization. Stakeholder input is critical in this phase, as their feedback ensures the system is user-friendly and tailored to meet the operational needs of law enforcement agencies. Iterative testing and refinement are essential, ensuring that the system evolves in response to real-world demands and feedback. The system is designed to be flexible, enabling it to adapt to changing crime trends and evolving user requirements.

Once the system is developed, the Compliance and Testing phase ensures that the system meets high ethical and legal standards. This includes conducting extensive validation tests on the predictive models for accuracy, fairness, and reliability, with a particular focus on avoiding biases that could impact the effectiveness of the system. The system undergoes rigorous compliance testing, ensuring it adheres to data protection laws, such as Malaysia's Data Protection Act, as well as international guidelines on the use of force and firearms. The system's performance is evaluated through a series of metrics, including precision, recall, and overall predictive accuracy, to ensure it can operate effectively in real-world law enforcement scenarios. In the System Deployment phase, the system is rolled out and integrated into operational environments. This stage includes comprehensive user training to ensure that law enforcement personnel are equipped with the necessary skills to use the system effectively. The deployment process is supported by on-going user engagement to guarantee that the system is both accessible and usable, enabling efficient integration into daily law enforcement operations. Finally, the Sustainability and innovation phase ensures that the system remains effective over time. This phase involves regular monitoring and evaluation of the system's performance, along with the identification and integration of innovations and improvements. The goal is to adapt the system to emerging crime trends, technological advancements, and evolving stakeholder needs, ensuring its long-term sustainability and relevance in enhancing crime prevention efforts. By following this structured approach, the research design ensures that the ICRAS system is developed, tested, and deployed in a way that is scientifically rigorous and practically impactful, capable of addressing the complex challenges of crime prevention in Malaysia.

Table 3.1
Phases of Crime Data Research: Quantitative and Qualitative Approaches

	Data Systems Architecture	System Design & Development	Compliance & Testing	System Deployment
Quantitative	Structuring databases, defining data formats, and implementing quantitative data models.	Coding algorithms, developing metrics, and building functionalities for the system.	Testing system performance, running regression tests, measuring accuracy rates, and ensuring adherence to benchmarks.	Monitoring system stability through metrics like downtime, response time, and error rates.
Qualitative	**Problem Definition** involves identifying the problem through exploratory methods such as literature review, expert interviews, stakeholder consultations, and thematic analysis to understand the scope and nature of the crime issues. **Knowledge Framework** Developing the knowledge framework involves understanding theoretical concepts, frameworks, and models through qualitative synthesis of academic literature, criminology theories, and policy analysis. **Data System Architecture** Understanding how data sources integrate, handling unstructured or semi-structured data, and assessing system requirements based on qualitative feedback. **System Design and Development** Incorporating user feedback into design prototypes, ensuring usability through interviews, and refining based on qualitative testing. **Compliance and Testing** ` Conducting user acceptance testing (UAT), collecting subjective feedback, and evaluating compliance with ethical guidelines. **System Deployment** Gathering user feedback on system functionality, interface satisfaction, and training effectiveness. **Sustainability and Innovation** evaluating the long-term adaptability of the system, exploring innovative enhancements, and ensuring alignment with emerging crime trends and stakeholder needs, primarily through qualitative insights.			

3.1.4 Integration of Data

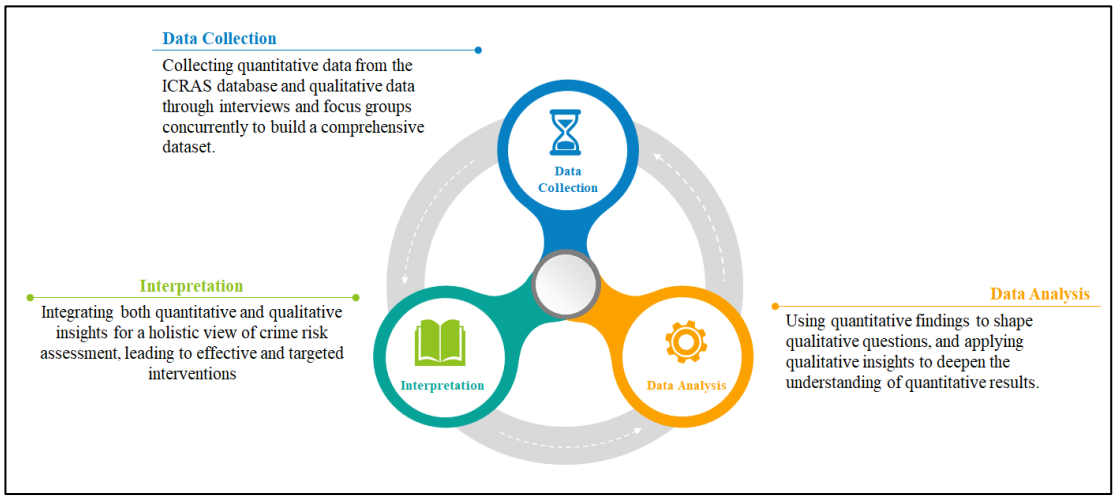


Figure 3.4 Integrating Quantitative and Qualitative Data for Comprehensive Crime Risk Assessment

i. Data Collection: Both types of data will be collected concurrently to ensure a comprehensive dataset. For example, while quantitative data is being gathered from the ICRAS database, qualitative data will be collected through interviews and focus groups.

ii. Data Analysis: Quantitative findings will inform the qualitative phase, guide the selection of interview questions and focus group discussions. For instance, if quantitative analysis reveals a spike in property crimes in a specific area, qualitative interviews with residents and law enforcement officers in that area can provide deeper insights into the underlying causes and potential solutions. Conversely, qualitative insights will provide context and deeper understanding to the quantitative results. For example, if interviews reveal that community members feel unsafe in poorly lit areas, this information can be used to guide quantitative analysis of crime rates in relation to street lighting .

iii. Interpretation: The final interpretation of the research findings will integrate insights from both datasets, providing a holistic view of the crime risk assessment system and its implications. For example, quantitative data may show a correlation between high unemployment rates and increased property crimes, while qualitative data may reveal that residents in high unemployment areas feel neglected by law enforcement. Integrating these insights can lead to more effective and targeted interventions .

The mixed methods approach is justified for this research due to its ability to address the complex and multifaceted nature of crime data, offering a comprehensive understanding those neither quantitative nor qualitative methods alone can provide. This approach encompasses several key benefits that significantly enhance the research process. Firstly, triangulation enhances the validity and reliability of the findings by cross-verifying data from multiple sources and methods. This process involves comparing quantitative data on crime rates with qualitative data from community interviews, ensuring a more accurate and comprehensive understanding of the crime situation. For instance, statistical data from crime reports can be juxtaposed with first hand accounts from residents and law enforcement officers to paint a fuller picture of the crime dynamics in a specific area . This cross-verification helps mitigate the biases and limitations inherent in each method, leading to more robust conclusions. Secondly, complementarity provides a more complete understanding by combining the numerical precision of quantitative data with the contextual richness of qualitative data. Quantitative data can reveal trends and patterns, such as a decrease in crime rates following the implementation of a new policing strategy. However, these numbers alone may not capture the community's perception of safety and the nuances of how the strategy is experienced by residents. Qualitative data fills this gap by providing insights into how the community perceives the new policing strategy and its impact on their sense of security and well being [176]. This complementary use of data enriches the overall understanding and allows for more informed decision-making. Furthermore, development is a key advantage of the mixed methods approach, as it involves using the results from one method to inform and refine the other, creating a more iterative and adaptive research process. Initial quantitative findings may highlight certain areas of concern that warrant further qualitative exploration. For example, if statistical analysis identifies a spike in property crimes in a particular neighbourhood, qualitative methods such as interviews and focus groups can delve into the underlying causes, such as social and economic factors, and suggest potential solutions. Conversely, insights gained from qualitative interviews may reveal new variables or hypotheses to be tested quantitatively, thus refining the scope and focus of subsequent quantitative analysis [177]. Additionally, expansion extends the breadth and scope of the research by exploring additional dimensions of the research problem. Qualitative data may uncover new factors influencing crime patterns that were not initially considered in the quantitative analysis. For instance, interviews with community members might reveal

that certain socio cultural dynamics, such as community cohesion or the presence of local leadership, play a significant role in crime prevention. This new information can lead to the inclusion of additional variables in the quantitative analysis, thereby broadening the research perspective and enriching the findings [178].

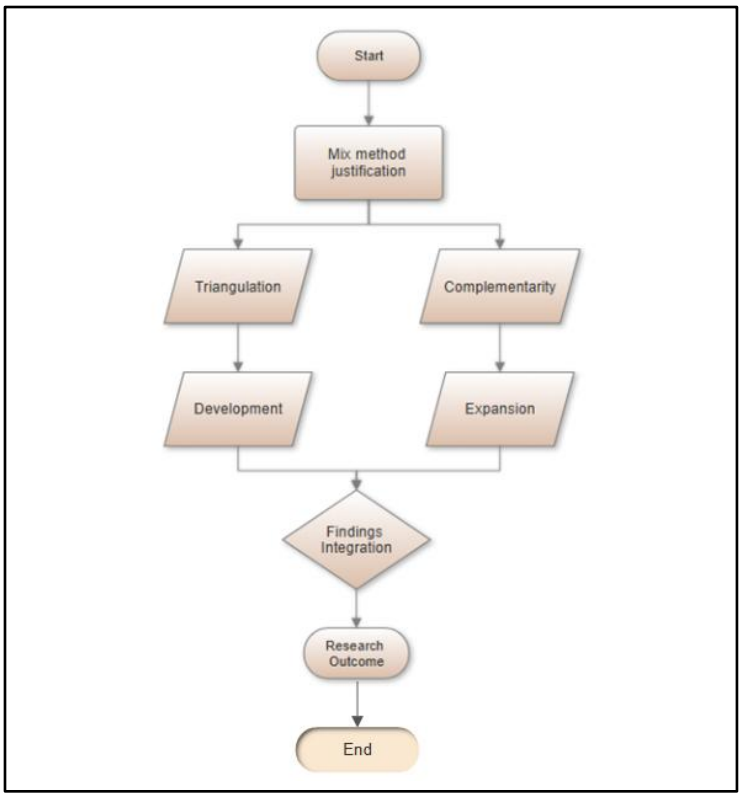


Figure 3.5 Mixed Methods Research Process

Implementing a mixed methods research design involves several challenges and considerations that need to be addressed to ensure the success of the study. One of the primary challenges is data integration, which requires effectively combining qualitative and quantitative data. This process demands careful planning and coordination to ensure that both types of data are collected, analysed, and interpreted in a coherent and complementary manner [179]. The integration of these data types is crucial for drawing comprehensive and insightful conclusions.

Another significant challenge is the resource intensity of mixed methods research. Collecting and analysing both qualitative and quantitative data requires considerable time, effort, and expertise. It is essential to allocate sufficient resources and ensure that the research team possesses the necessary skills and experience to handle both types of data effectively [180]. This often means having a team with diverse

skill sets and ensuring adequate training and support. Ensuring the quality and consistency of data is also critical. This involves implementing rigorous data pre-processing and cleaning procedures for quantitative data, as well as meticulous transcription and coding processes for qualitative data [181]. High-quality data are essential for producing reliable and valid research findings and maintaining consistency across data types is crucial for meaningful integration and analysis. Ethical considerations are paramount in mixed methods research, especially when dealing with sensitive topics such as crime. Researchers must ensure informed consent, protect the confidentiality and privacy of participants, and address any potential risks or harms associated with the research [182]. Ethical considerations also include being transparent with participants about the study's aims and ensuring their voluntary participation. Stakeholder engagement throughout the research process is crucial for ensuring the relevance and impact of the findings. Collaborating with law enforcement agencies, community organizations, and other relevant stakeholders ensures that their perspectives and insights are considered [183]. This collaboration helps to ground the research in real-world contexts and enhances its applicability and usefulness.

Data interpretation presents another challenge, as it involves integrating findings from both types of data, which can be complex and may require iterative analysis and validation. Ensuring that the interpretations are coherent and meaningful demands careful attention to detail and a thorough understanding of both quantitative and qualitative methods [184]. This process often involves revisiting the data multiple times and refining the interpretations based on emerging insights. Interdisciplinary collaboration is often necessary in mixed methods research, requiring teamwork among researchers with different areas of expertise. This interdisciplinary approach can enhance the study by incorporating diverse perspectives and skills, but it also necessitates effective communication and coordination to align the different methodological approaches. Successful collaboration involves fostering an environment of mutual respect and open communication. Finally, managing large datasets is a significant challenge in mixed methods research. Handling extensive quantitative datasets and qualitative data requires efficient data management practices, including the use of appropriate software tools for data storage, organization, and analysis [185]. Proper data management ensures that data are accessible, secure, and analysable, facilitating the overall research process and enhancing the quality of the findings. While mixed-methods research offers comprehensive insights and robust

findings, it also presents several challenges that require careful consideration and meticulous planning. Addressing these challenges effectively involves integrating data, managing resources, and ensuring data quality, upholding ethical standards, engaging stakeholders, interpreting data accurately, fostering interdisciplinary collaboration, and managing large datasets efficiently. By tackling these challenges, researchers can maximize the potential of mixed methods research to produce valuable and impactful results.

3.2 Population and Sampling

3.2.1 Definition of Population

The population for this study comprises all recorded crime incidents within the Shah Alam district from the years 2015 to 2021. Shah Alam, located in the state of Selangor, Malaysia, is a prominent urban centre with a diverse population and economic activities ranging from residential areas to industrial zones. Crime incidents within this timeframe are systematically documented by law enforcement agencies, particularly the Shah Alam Police Department, and include a variety of criminal offenses categorized under property crimes (e.g., theft, burglary) and violent crimes (e.g., assault, robbery) . Shah Alam's demographic and socioeconomic diversity influences crime dynamics, making it a crucial area for studying crime trends and patterns. The population of interest encompasses all reported crime incidents within the district, offering a comprehensive dataset for analysing the spatial, temporal, and situational aspects of crime over more than a decade.

The sampling strategy for this research adopts a systematic approach to ensure the selection of a representative sample of crime incidents from the larger population within the specified timeframe. This approach addresses practical constraints associated with extensive data collection while maintaining the reliability and generalizability of findings to the broader population. In criminological research, the importance of a robust sampling strategy cannot be overstated, as it underpins the validity of subsequent analyses and conclusions [186]. For this study, which investigates crime incidents in Shah Alam from 2010 to 2021, a systematic sampling method was identified as the most suitable. Systematic sampling involves selecting every k-th crime incident from an ordered list of records, ensuring that all cases have an equal probability of inclusion.

This method not only enhances representativeness but also minimizes sampling bias by maintaining the integrity of the dataset structure. For example, in a study on urban crime trends in Kuala Lumpur, systematic sampling was employed to analyse theft cases over a decade, demonstrating the method's utility in generating reliable insights while managing voluminous datasets [187]. In this research, the sequential ordering of crime incidents was based on temporal and spatial criteria. For instance, within each stratum defined by year and geographic district every 10th crime incident was selected for analysis. This approach aligns with prior studies, such as those by [188], which employed systematic sampling to investigate patterns in burglary incidents across metropolitan areas, and [189], who used similar techniques to evaluate crime deterrence strategies in high density neighbourhoods. The choice of systematic sampling over other methods, such as random or stratified sampling, was informed by its practicality and its ability to maintain the chronological and spatial distribution of crime incidents. This methodological consideration ensures that key trends, such as seasonal spikes in crime or location specific concentrations, are captured accurately. Additionally, systematic sampling has been shown to reduce data redundancy and processing time, as highlighted in comparative studies by [190].

Overall, the systematic sampling strategy employed in this research adheres to established best practices and methodological standards in criminology, leveraging its strengths in achieving both efficiency and representativeness. The use of systematic sampling facilitates robust analysis while addressing the challenges posed by large and diverse crime datasets, as demonstrated in similar studies across urban settings [191].

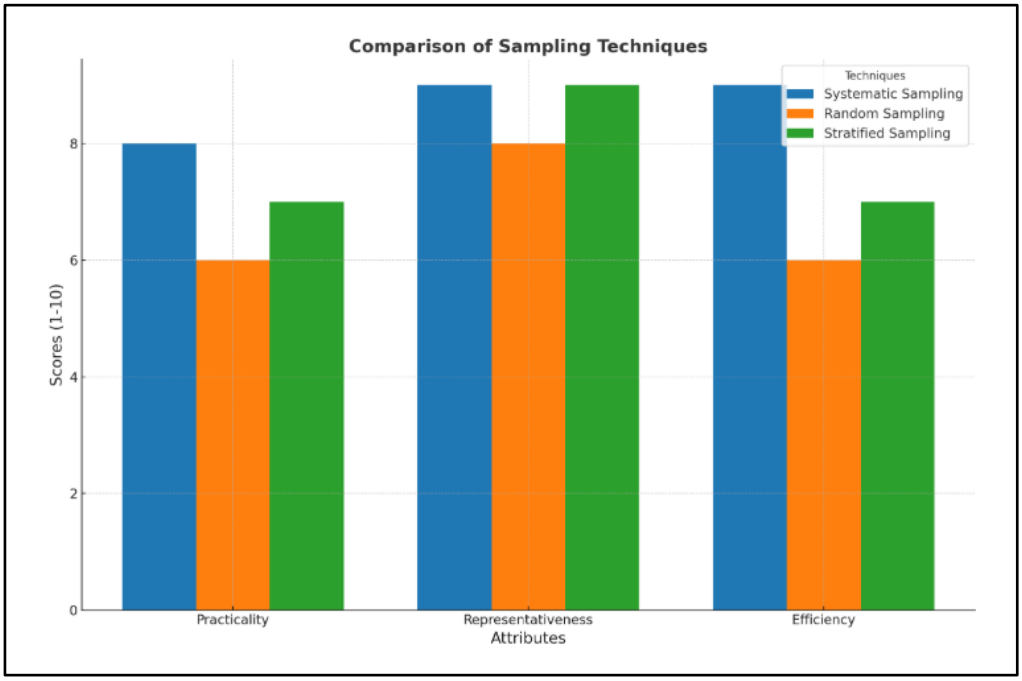


Figure 3.6 Comparison of Sampling Techniques

The comparison chart illustrates the evaluation of three widely recognized sampling techniques Systematic Sampling, Random Sampling, and Stratified Sampling across the dimensions of practicality, representativeness, and efficiency. These dimensions are critical in determining the suitability of a sampling method for criminological research, particularly when handling large and diverse datasets. Systematic Sampling, represented as the highest-performing technique, demonstrates a compelling balance between efficiency and representativeness. Its efficiency score of nine (9) reflects its robustness in processing voluminous datasets while minimizing redundancy, as highlighted in studies such as [192]. Moreover, its high representativeness score of nine (9) underscores its ability to capture the underlying population's diversity without introducing significant bias, as discussed in [193] exploration of crime deterrence strategies. While its practicality is slightly lower, scoring 8, it remains a highly feasible choice, especially in contexts where ordered datasets are readily available, aligning with findings from [194] on urban crime trends in Kuala Lumpur. In contrast, Random Sampling, despite its methodological simplicity, scores comparatively lower across all dimensions. This technique’s practicality score of 6 reflects its limitations in real-world applications, where true randomness often requires specialized tools or algorithms, as noted by [195]. The representativeness score of eight (8) suggests that while randomness ensures fairness in selection, it can result in gaps in the dataset, potentially missing critical trends or patterns. Its efficiency score of

6 highlights challenges related to data redundancy and the lack of inherent order, a concern echoed in comparative analyses by [196] These limitations position random sampling as less optimal for studies requiring nuanced insights into spatiotemporal trends, such as crime seasonality or geographic hotspots. Stratified Sampling, with its emphasis on representativeness, achieves the highest score of 9 in this dimension by ensuring proportional inclusion of all subgroups within the population. This aligns with findings from [197], who emphasized the technique's capacity to reflect demographic or geographic heterogeneity in crime data. However, the practicality score of 7 highlights the complexity involved in defining and categorizing strata, which can be resource intensive, as observed in [198] studies on metropolitan burglary patterns. Efficiency, while moderately high at 7, also reflects the balance this method strikes between comprehensive data coverage and feasibility, as discussed by [199]. The chart ultimately underscores the superiority of systematic sampling for the present research, given its strong performance across all evaluated dimensions. By maintaining the integrity of the dataset's temporal and spatial order, systematic sampling captures critical trends like seasonal spikes in crime and location specific concentrations, which are pivotal in criminological analyses. This aligns with prior studies such as those by [200] and [201], which have validated the method’s utility in crime related research. The chart further highlights the inherent trade-offs among the techniques, offering a nuanced perspective on their relative strengths and limitations. By leveraging insights from multiple empirical studies, this comparison substantiates the rationale for selecting systematic sampling, demonstrating its alignment with established best practices in criminology while ensuring efficiency and representativeness in large scale data analysis.

Table 3.2
Temporal Stratification of Crime Incidents by Year (2015–2021)

Stratification by Crime Type and Year	Temporal Stratification
Crime incidents are categorized into two main types property crimes (including theft, burglary, and vandalism) and violent crimes (such as assault, robbery, and homicide). Each crime type exhibits distinct patterns and impacts on community safety, necessitating separate analysis to understand their dynamics comprehensively.	The study spans seven years from 2015 to 2021, encompassing various socio political and economic changes that may influence crime trends. Temporal stratification ensures that the sample reflects changes in crime patterns over time, capturing both short-term fluctuations and long-term trends in crime incidence within Shah Alam [121]

To ensure a comprehensive analysis of crime patterns, this study employs a stratified approach based on both crime type and temporal factors. Crime incidents are categorized into two main types: property crimes, which include theft, burglary, and vandalism, and violent crimes, which encompass offenses such as assault, robbery, and homicide. Each category is analyzed separately to capture distinct patterns and their impact on community safety, as different crime types exhibit unique characteristics and influencing factors. By classifying crimes in this manner, the study ensures that variations in their occurrence, severity, and distribution are thoroughly examined, providing deeper insights into their dynamics within Shah Alam. In addition to crime type stratification, temporal stratification is incorporated to account for both short-term fluctuations and long-term trends in crime incidence. The study spans a seven year period from 2015 to 2021, a timeframe that encompasses various socio political and economic changes that may have influenced crime trends. By structuring the dataset across multiple years, this approach enables the identification of patterns that emerge over time, allowing for a clear understanding of how crime rates evolve in response to external factors. This ensures that the analysis does not merely capture isolated incidents but rather reflects broader changes in crime behavior, law enforcement strategies, and societal conditions. Through this dual stratification, the study provides a more robust and detailed examination of crime trends, enhancing its ability to generate meaningful insights that contribute to crime prevention and law enforcement efforts.

3.2.2 Data Collection Procedures

Data collection involves accessing and compiling crime incident records from official sources, primarily the Shah Alam Police Department and other relevant government agencies responsible for crime documentation and reporting. The data collection process adheres to ethical guidelines, ensuring confidentiality, privacy, and responsible use of sensitive information related to crime incidents and individuals involved [202].

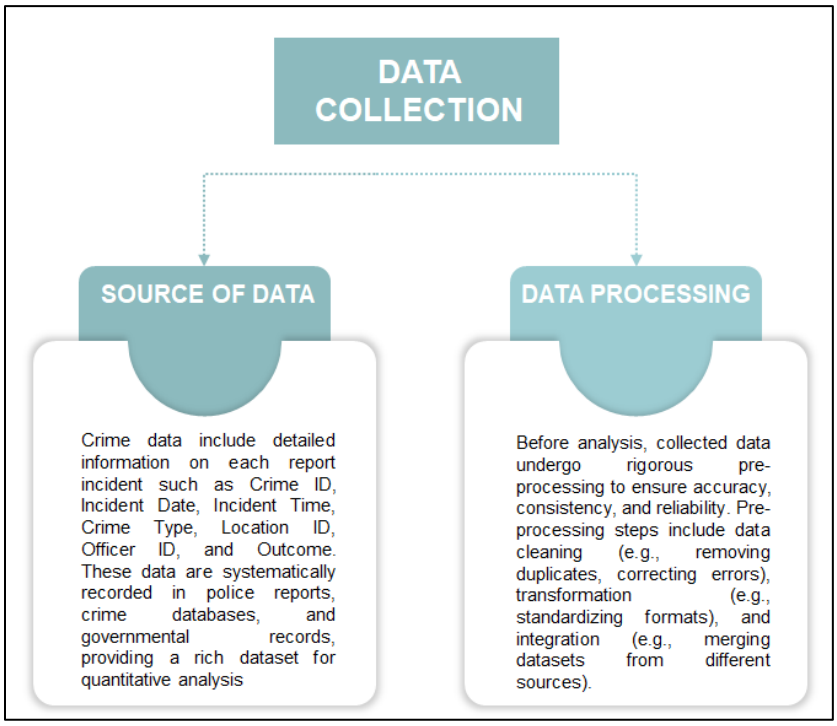


Figure 3.7 Overview of Crime Data Collection and Pre Processing Methodology

3.2.3 Sampling Procedure

For this study, the sampling procedure was carried out using systematic sampling, a probability sampling method that involves selecting participants at regular intervals from an ordered list. This technique was deemed most appropriate for ensuring a well structured and representative sample of participants, drawn from key stakeholders involved in crime data analysis, risk assessment, and system administration. The systematic sampling approach ensured that each participant had an equal opportunity to be selected, thus reducing the potential for selection bias and ensuring diversity within the sample. The participants for the evaluation of the Integrated Crime Risk Assessment System (ICRAS) were selected from three distinct stakeholder groups within the Malaysian law enforcement community. These groups included Investigation Officers (IOs) from the Royal Malaysian Police (RMP), crime analysts, and information technology personnel responsible for managing and maintaining crime reporting systems and databases. These categories were identified as the most pertinent for assessing ICRAS, as they each contribute essential expertise to the system's operation, analysis, and technical functionality. A total of thirty (30) participants were selected for the pilot evaluation, with an equal representation from each of the three stakeholder groups. The participants were systematically chosen from

an ordered list of eligible individuals provided by the RMP, crime analysis departments, and the IT division. This approach ensured that the sample reflected a broad range of perspectives, with insights drawn from operational, analytical, and technical viewpoints. The systematic selection of participants allowed for an even distribution of input from those directly involved in the daily use and management of crime data systems, enhancing the comprehensiveness of the system evaluation. The decision to use systematic sampling aligns with the research methodology for this study, which follows a mixed-methods approach. This methodology combines qualitative and quantitative techniques to evaluate the functionality, usability, and predictive accuracy of ICRAS. By focusing on a structured and objective sampling method, systematic sampling provided a solid foundation for obtaining meaningful and representative feedback from participants, who were selected based on their expertise and direct experience with crime data systems.

It is important to note the distinction between the System Development Methodology and the Research Methodology in this study. The System Development Methodology, such as Agile or Software Development Life Cycle (SDLC), pertains to the process of designing and building the ICRAS system, while the Research Methodology focuses on the evaluation process of the system through participant feedback. The systematic sampling technique specifically applies to the evaluation phase of this research, which is distinct from the development process and is critical for obtaining actionable insights on ICRAS's effectiveness and practical application. By adopting systematic sampling, the study ensured that the evaluation of ICRAS was informed by participants who possess specialized knowledge and relevant experience, ultimately enhancing the reliability and depth of the findings. This approach not only supports the validity of the study but also ensures that the feedback collected is reflective of authentic operational realities and provides valuable implications for the future scalability and enhancement of ICRAS across Malaysia.

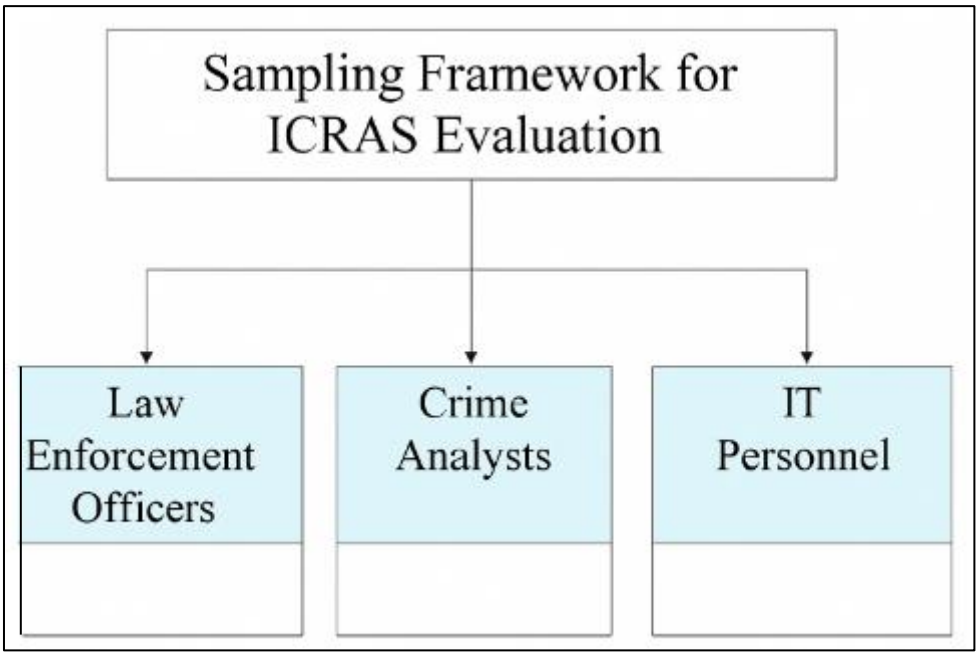


Figure 3.8 Systematic Sampling Framework for ICRAS Evaluation

3.2.4 Data Processing Using the Extract, Transform, and Load (ETL) Framework

To preserve the integrity, consistency, and usability of the datasets used in this study, an Extract, Transform, and Load (ETL) process was applied as the central data processing mechanism. This step was crucial for managing the diversity and complexity of crime data obtained from several sources, including Royal Malaysian Police (RMP) records, statistical reports, and other related repositories. The ETL process ensured that all information was standardized, cleansed, and structured in a form suitable for reliable analysis within the Integrated Crime Risk Assessment System (ICRAS). The first stage, extraction, involved retrieving data from various independent systems such as RMP incident logs, crime statistics, and other law enforcement databases. Because these datasets differed in format and structure some being structured tables while others were text-based or semi-structured a systematic extraction strategy was necessary. During this process, key variables such as crime type, location, date, time, and demographic attributes were identified and captured. Particular care was taken to uphold data confidentiality and privacy throughout, ensuring that sensitive records were handled securely while maintaining the completeness of analytical variables. Once data from multiple sources had been consolidated, it proceeded to the transformation phase. This stage was aimed at improving quality and analytical readiness. The process began with

data cleaning, which involved identifying and correcting missing values, duplicates, and outliers. Errors such as inconsistent date formats or misclassified crime categories were also rectified. These steps were critical for ensuring that subsequent analyses were based on accurate and trustworthy data. Following the cleaning phase, the data was normalized to ensure consistency across variables and datasets. Temporal and spatial fields were standardized to enable meaningful comparisons across regions and time periods, while categorical variables were recoded into uniform classifications. In addition, supplementary demographic and socio-economic data were integrated to enrich the dataset, thereby improving the predictive accuracy of the ICRAS model. The final stage, loading, involved transferring the processed data into a centralized repository designed for efficient access, real-time analysis, and visualization. This structured repository served as the backbone for statistical and risk assessment processes, enabling law enforcement officers and analysts to access consolidated, high-quality information in support of investigations and strategic decision making.

Through the implementation of this ETL framework, the study ensured that all analytical outcomes were built upon reliable, well-integrated, and high-quality data. In essence, ETL provided the technical bridge between raw data acquisition and the empirical crime analysis that follows. By transforming fragmented inputs into a coherent analytical foundation, this process strengthened both the validity and operational value of the ICRAS model. The datasets prepared through ETL were subsequently used for statistical modeling and pattern identification, as detailed in Section 3.2.6.

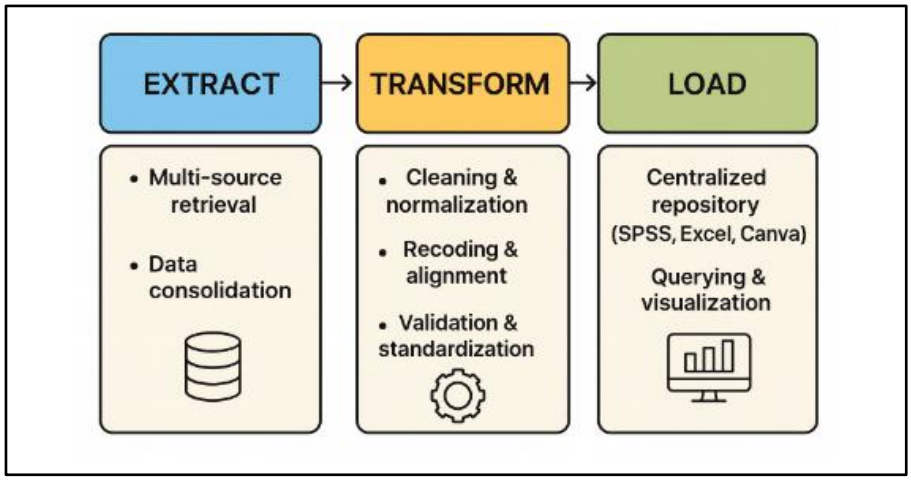


Figure 3.9 ETL Based Data Processing Workflow for ICRAS

3.2.5 Data Analysis and Interpretation

Upon completion of data collection and pre-processing, the sampled crime data undergo comprehensive quantitative analysis to achieve research objectives and address research questions related to crime patterns, trends, and impacts within Shah Alam. Statistical analysis techniques, including descriptive statistics, inferential statistics and temporal analysis are applied to derive actionable insights and evidence based recommendations [203].

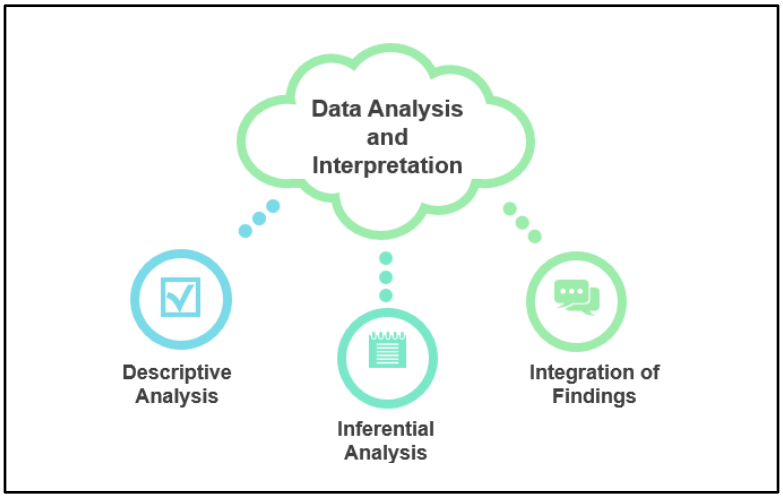


Figure 3.10 Comprehensive Frameworks for Data Analysis and Interpretation

i. Descriptive analysis summarizes key characteristics and distributions of crime incidents within each stratum, providing an overview of crime patterns by type and year. Measures such as frequencies, percentages, central tendency (mean, median), and dispersion (standard deviation) are calculated to describe the prevalence and variability of crime incidents over time [204].
ii. Inferential analysis tests hypotheses and examines relationships between variables within the sampled data. Techniques such as correlation analysis, regression analysis, and hypothesis testing evaluate statistical associations and identify predictors of crime incidence, law enforcement effectiveness, and community safety outcomes. Inferential statistics provide insights into factors influencing crime dynamics and inform strategic interventions to mitigate crime risks within Shah Alam [205].
iii. Findings from quantitative analysis are integrated with qualitative insights obtained from stakeholder interviews and focus groups (discussed in Chapter

3.3) to provide a holistic understanding of crime dynamics and community perceptions within Shah Alam. Integrated analysis supports evidence based decision making, policy formulation, and targeted interventions aimed at enhancing public safety and reducing crime rates in urban environments.

3.2.6 Methodological Considerations and Challenges

Implementing the sampling strategy involves addressing several methodological considerations and challenges to ensure the reliability and validity of study findings:

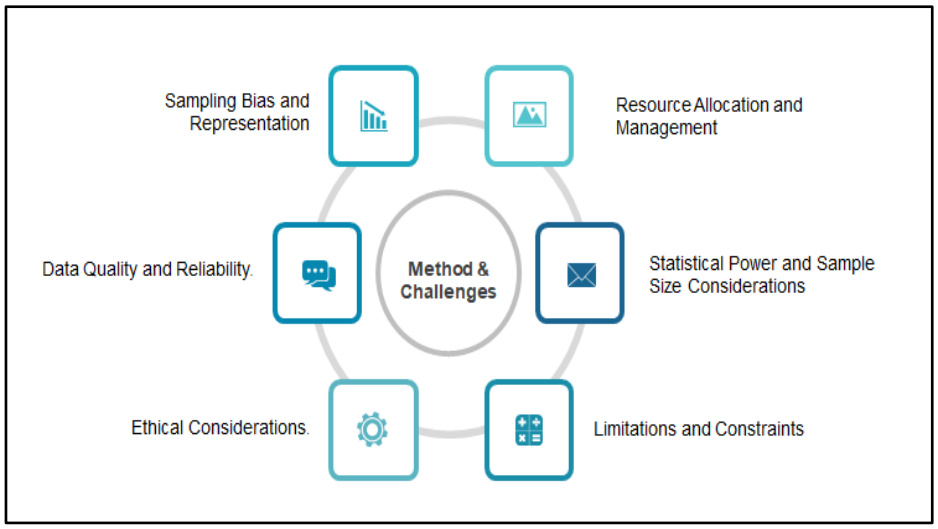


Figure 3.11 Methodological Challenges and Key Considerations

a) Sampling Bias and Representation: Efforts are made to minimize sampling bias by adhering to predefined sampling criteria, ensuring transparency in the selection process, and maintaining proportionate representation of crime types and temporal periods within the sample. The systematic sampling procedure mitigates biases associated with non-random selection and enhances the representativeness of the sampled data [206].

b) Data Quality and Reliability: Ensuring the accuracy, consistency, and reliability of crime data is paramount to the integrity of statistical analyses and research conclusions. Data validation procedures, including cross-checking with multiple sources and verifying against established criteria, are employed to enhance data quality and minimize errors in recorded crime incidents [207].

c) Resource Allocation and Management: Allocating resources effectively—such as time, funding, and personnel is essential to the successful implementation of

sampling procedures and data collection activities. Adequate planning, coordination, and supervision of research activities ensure timely acquisition of data and adherence to project timelines, enhancing the efficiency and productivity of study operations [208].

d) Statistical Power and Sample Size Considerations: Determining an appropriate sample size and ensuring sufficient statistical power are critical to the validity and reliability of statistical analyses. Statistical formulas and calculations guide sample size determination, considering variability in crime incidents, desired level of confidence, and acceptable margin of error. A sufficiently large sample size enhances the precision of estimates and strengthens the validity of research findings, supporting robust statistical inference and meaningful interpretation of results [209].

e) Limitations and Constraints: Despite rigorous methodological approaches, the sampling strategy may encounter limitations and constraints inherent to research design, data availability, and practical considerations. Acknowledging and addressing these limitations such as potential gaps in data coverage, challenges in accessing historical records and resource constraints contributes to a transparent and comprehensive discussion of the study's scope and implications.

3.3 Quantitative Data Collection

Quantitative data is indispensable for understanding crime trends in Shah Alam, as it provides the backbone for statistical analysis, revealing objective patterns across a large dataset. By using measurable variables and large scale crime data, quantitative methods allow the researcher to generate precise conclusions regarding crime distribution, frequency, and temporal trends. The study applies this method to analyse the evolution of crime incidents between 2015 and 2021, and to explore correlations between types of crimes, locations, and socio economic factors.

3.3.1 Sources of Quantitative Data

The quantitative data for this study was obtained from multiple official and authoritative sources, ensuring accuracy and reliability. The primary source of data was the Shah Alam Police Department, which holds detailed records of crime incidents

within the district. These data were complemented by additional resources from national crime statistics reports and publicly available government databases.

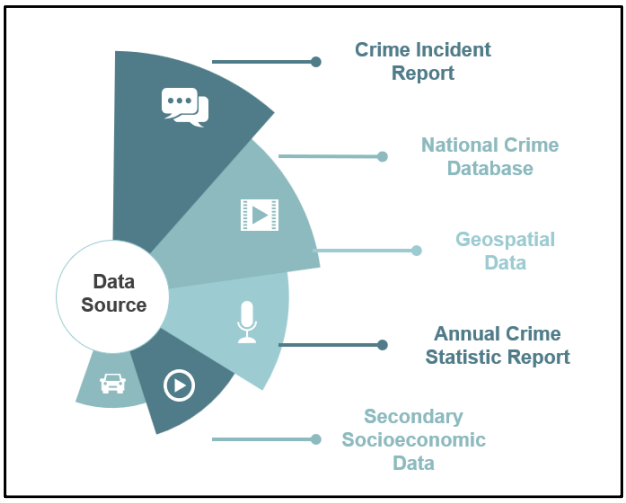


Figure 3.12 Key Data Sources for Crime Analysis in ICRAS

3.3.2 Crime Data Extraction, Cleaning, and Processing

The process of collecting quantitative data followed a structured methodology designed to ensure completeness, accuracy, and ethical compliance. The data collection was executed in several stages, each aimed at ensuring that the most relevant and reliable information was gathered for analysis.

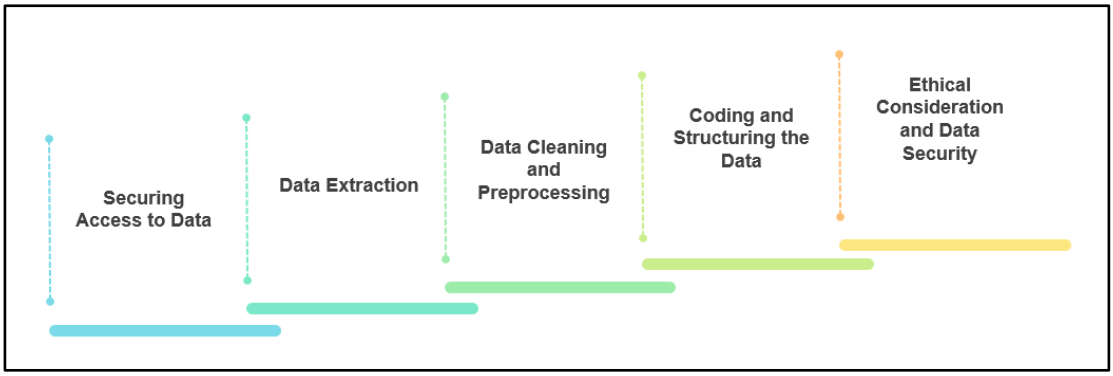


Figure 3.13 Data Processing and Ethical Workflow for ICRAS Development

1. Securing Access to Data: Gaining access to police and governmental data required formal procedures. A request was submitted to the Shah Alam Police Department for access to crime records, supported by an ethics clearance certificate from the overseeing academic institution. This formal agreement

guaranteed that the data would be used solely for research purposes, in line with ethical guidelines that protect both the privacy of individuals involved and the security of sensitive crime data. Additionally, the research team obtained permission from national bodies such as the Royal Malaysia Police and the Department of Statistics Malaysia to access broader databases. Data use agreements were signed, outlining the conditions under which data could be analysed and reported.

2. Data Extraction**:** Upon gaining access, the research team worked closely with law enforcement agencies to extract relevant crime data from their internal systems. Given the sensitivity of the data, this extraction was conducted in a secure environment, with crime reports and datasets transferred digitally in encrypted formats. The data extraction process focused on selecting cases that met specific criteria, including:
    i. All crime incidents reported in Shah Alam from 2015 to 2021.
    ii. Inclusion of both property crimes (e.g., burglary, theft) and violent crimes (e.g., assault, robbery).
    iii. Detailed records with complete temporal and geographic information.

The goal of the extraction was to create a large, longitudinal dataset that could be systematically analysed for trends over the 7 year period .

3. Data Cleaning and Pre-processing: The raw data obtained from police records often contained inconsistencies and missing information. As a result, the first step in data preparation was cleaning the dataset. This involved:
    i. Handling Missing Data: Crime reports with incomplete or unclear information were lagged, and either corrected (where supplementary information was available) or removed from the final dataset.
    ii. Correcting Errors: Minor errors, such as typographical mistakes in crime categories or dates, were corrected by cross-referencing with official reports and annual crime statistics.
    iii. Standardizing Variables: To ensure compatibility across datasets, variables such as crime types, location identifiers, and dates were standardized into a consistent format. For example, crime types were classified into broader

categories (property vs. violent crime), and location data was aligned with GIS coordinates for later spatial analysis.

iv. Removing Duplicates: In cases where duplicate reports for the same incident existed (e.g., multiple officers filing reports on the same crime), these duplicates were identified and removed to avoid skewing the analysis.

3.3.3 Quantitative Data Collection

While quantitative data provides an objective, statistical understanding of crime patterns in Shah Alam, qualitative data offers a more subjective but deeply contextual perspective. Through qualitative methods, the study gathers insights from key stakeholders such as law enforcement officers, community members, and victims of crime. These perspectives contribute to a fuller understanding of the societal impacts of crime, as well as the effectiveness of crime prevention strategies employed in the district. The qualitative approach is designed to complement quantitative findings by addressing the "why" and "how" behind the observed crime patterns. It explores the subjective experiences of individuals affected by crime and those involved in crime prevention. For instance, while quantitative data can reveal an increase in crime rates, qualitative data can provide insights into community perceptions of safety and law enforcement responses. Qualitative data was collected through several key methods, including in-depth interviews, focus group discussions, and content analysis of crime-related media coverage. Each method provided a different lens through which to view the crime issues in Shah Alam, ensuring the data captured was comprehensive and multifaceted.

a. In Depth Interviews - In-depth, semi-structured interviews were conducted with three primary groups:
    i. Law Enforcement Officers: These interviews focused on understanding the practical challenges officers face in preventing and addressing crime. Topics included resource limitations, operational strategies, and the officers' perspectives on crime trends.
    ii. Community Leaders: As representatives of the broader population, community leaders provided valuable insights into how local communities perceive crime. Their interviews highlighted community concerns, the

perceived effectiveness of law enforcement, and the social factors contributing to crime.

iii. Victims of Crime: Interviews with victims offered personal insights into the consequences of crime. These discussions delved into emotional and psychological impacts, perceptions of safety, and experiences with the criminal justice system .

iv. Resident: Interviews with residents from both high-crime and low-crime areas provided valuable insights into their experiences with safety and crime prevention. These discussions explored their perceptions of neighbourhood security, interactions with law enforcement, and community led initiatives to reduce crime. Additionally, the study examined differences in fear of crime, trust in the justice system, and the role of social cohesion in shaping crime prevention efforts.

Each interview was conducted using a semi structured approach, with a core set of questions but enough flexibility to allow interviewees to express their experiences freely. This method helped ensure that important, unforeseen topics could emerge naturally from the conversations.

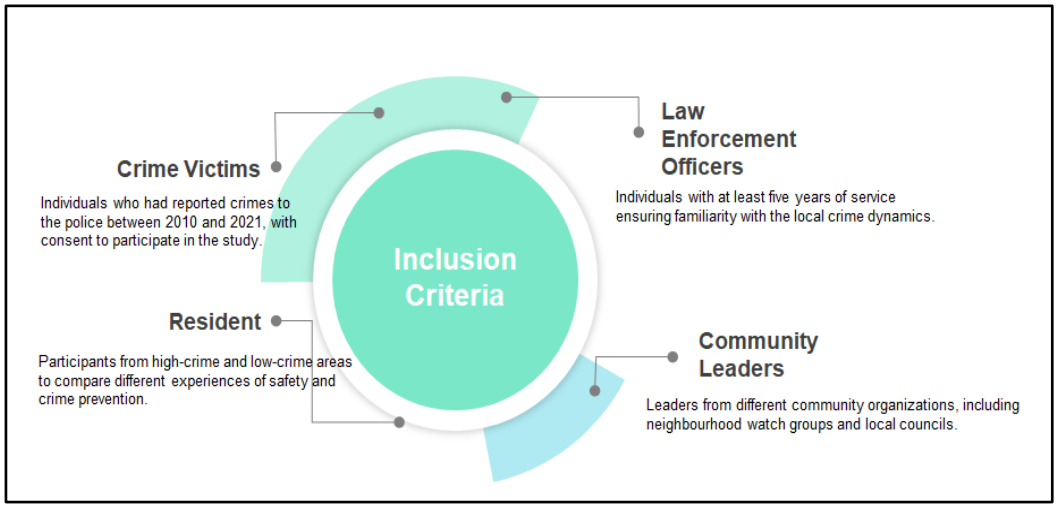


Figure 3.14 Inclusion Criteria for Stakeholder Participation in the ICRAS Study

3.3.4 Data Management Systems

The research collected large volumes of data over a period of seven years (2015–2021) from various sources, including police records, interviews, focus group discussions, and public reports. To ensure the data was well organized, secure, and easily accessible for analysis, a systematic approach to data management was adopted. The following subsections outline the systems and procedures implemented for managing both quantitative and qualitative data

i. Quantitative Data Management

The quantitative data in this study primarily involved crime statistics from Shah Alam, sourced from police databases and government reports. Properly managing this data required not only secure storage but also careful pre processing to ensure its accuracy and usability in statistical analyses.

a) Data Storage: All quantitative data was stored in a secure, cloud-based research data management system (RDMS). This system allowed for centralized storage of large datasets, ensuring easy access for authorized personnel while maintaining strict data security protocols. Data backups were conducted daily, and all files were encrypted to protect against unauthorized access.
b) File Formats: The crime data was stored in standard formats such as CSV (Comma-Separated Values) and Excel spread sheets. These formats facilitated ease of use in statistical software (e.g., SPSS, R) and ensured compatibility with data pre-processing tools.

Data Pre processing - Before analysis, quantitative data underwent several pre-processing steps:

a) Data Cleaning: This involved identifying and removing duplicate entries, correcting erroneous values, and dealing with missing data. For example, if a crime incident was logged with incomplete details, it was cross referenced with official police records to fill in gaps or confirm its accuracy.

b) Data Normalization: Crime data from different years and police jurisdictions sometimes used varying formats and terminologies. For instance, a burglary might have been classified differently in earlier years. Data normalization ensured that such inconsistencies were harmonized, allowing for more accurate comparative analysis.
c) De-identification: To safeguard the identities of individuals in crime reports, all personal details (such as names and addresses) were removed or masked. This measure was essential to meet ethical standards and comply with data protection laws.
d) Data Categorization - Once cleaned and normalized, the quantitative data was categorized into different crime types (e.g., property crimes, violent crimes) and temporal periods (e.g., yearly breakdowns). These categorizations enabled more granular analysis and easier cross-referencing with qualitative data later in the study.

ii. Qualitative Data Management

Qualitative data management presented its own challenges, particularly due to the more subjective and narrative nature of the data. This section outlines how interview and focuses group transcripts, as well as media content, were systematically stored and organized for analysis.

Data Recording and Storage - All qualitative data, including interview and focus group transcripts, were stored in the same secure RDMS as the quantitative data. Audio recordings of interviews and discussions were first transcribed into text, after which the original audio files were securely archived for future reference if needed.

a. File Formats: Transcripts were stored in text and Word formats, allowing for easy coding and thematic analysis. Audio files were stored in MP3 format, ensuring compatibility with transcription software.

Data Organization - To facilitate the analysis process, qualitative data was organized into a database with key metadata fields, such as the date of the interview, the type of participant (law enforcement officer, community leader, victim, etc.), and

major themes covered in the conversation. This metadata categorization allowed for easy retrieval of specific interviews or discussions relevant to research questions.

Transcription Accuracy and Validation - To ensure the accuracy of the transcriptions, a two step validation process was implemented. First, the initial transcripts were generated using speech-to-text software. Then, human review was conducted by trained research assistants who listened to the original audio and corrected any misinterpretations made by the software. This method ensured that the final transcripts accurately reflected the content of the interviews and discussions.

Ethical Considerations in Storage - Given the sensitive nature of the topics discussed, especially during interviews with crime victims, strict confidentiality measures were applied. All qualitative data files had identifying information removed, and any personal details about participants were omitted or concealed in the transcripts. Access to these data was tightly restricted, with only authorized researchers allowed to view complete transcripts.

3.3.5 Key Components of the Consent Process

i. Purpose of the Study: Participants were informed that the research sought to understand crime patterns in Shah Alam, with a focus on the impact of crime on the community and the effectiveness of law enforcement efforts.
ii. Use of Data: Participants were made aware that their interviews would be recorded, transcribed, and analysed for research purposes. They were assured that their names and any other identifying information would be anonymized in the final report.
iii. Voluntary Participation: Participants were explicitly told that their involvement in the research was entirely voluntary, and that they could choose to withdraw from the study at any point, without having to provide a reason.
iv. Confidentiality: It was made clear that all data would be kept confidential, and that identifying information would be removed to protect participants' privacy.

3.3.6 Privacy and Anonymity

The anonymity of participants was a paramount concern, particularly for crime victims who had already experienced trauma. During the data collection process, every effort was made to protect their identities. This was achieved by assigning each participant a unique identifier (e.g., P1, FG1) and removing any personal details from the transcripts. In addition to removing personal identifiers from the data, strict confidentiality protocols were established. Access to the raw data was limited to the research team, and no information was shared with external parties without obtaining explicit consent from the participants. Furthermore, the data was stored in encrypted formats to prevent unauthorized access.

Handling Sensitive Data

Much of the qualitative data collected in this study involved sensitive discussions about personal experiences with crime. To ensure that these interviews were conducted ethically, interviewers were trained to adopt a trauma-informed approach. This involved recognizing the potential emotional effects of discussing past trauma and taking steps to reduce harm.

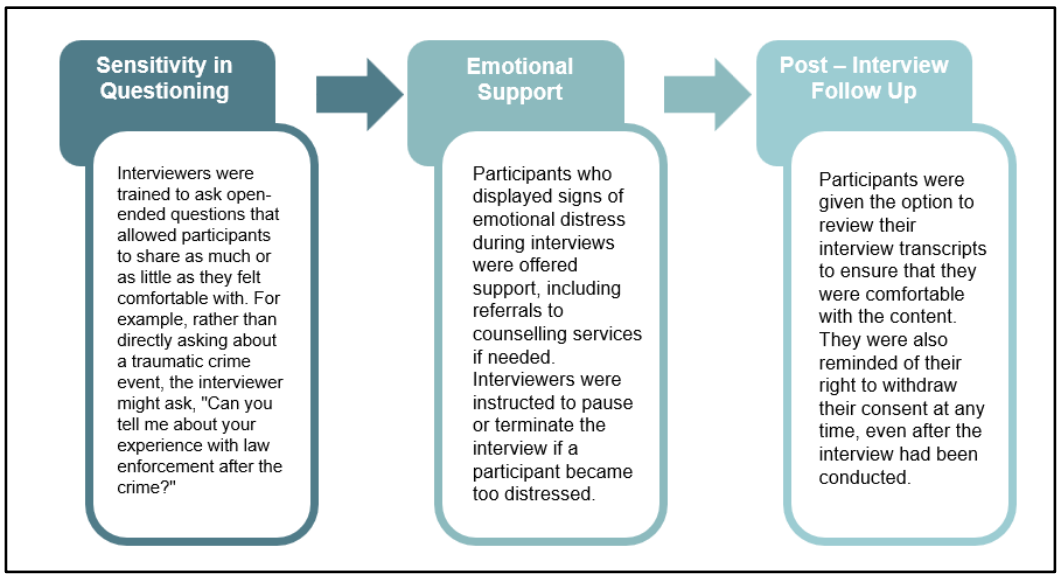


Figure 3.15 Trauma Informed Techniques

3.3.7 Compliance with Data Protection Laws

The study fully complied with all relevant data protection laws, particularly Malaysia's Personal Data Protection Act (PDPA), which governs the collection, use, and storage of personal data. Given the sensitive nature of crime-related data, a comprehensive framework was implemented to uphold ethical and legal standards while ensuring the confidentiality and security of participants' information. A key pillar of this compliance was the principle of informed consent, whereby all participants were explicitly informed of the study's purpose, the type of data being collected, and its intended use. This transparency not only fulfilled legal obligations but also fostered trust, ensuring that individuals felt secure in their participation. Moreover, participants were granted full autonomy over their involvement, including the right to withdraw their consent at any stage without consequence. To further safeguard privacy, the study strictly adhered to the principle of data minimization, ensuring that only essential personal information was collected. Unnecessary data fields were meticulously excluded to maintain a focused and relevant dataset, thereby reducing the risk of exposure to sensitive information. This deliberate limitation of data collection aligned with both ethical research standards and regulatory requirements, demonstrating a commitment to responsible data handling. In addition to minimizing data collection, anonymization techniques were employed to remove personally identifiable information wherever possible. By de-identifying sensitive data, the study significantly reduced the likelihood of individuals being traced through the collected information, reinforcing the integrity of the research process. Recognizing the paramount importance of data security, advanced encryption protocols were implemented to protect all digital records from unauthorized access. Secure encryption mechanisms, including Advanced Encryption Standard (AES) and Transport Layer Security (TLS), were employed to safeguard data at rest and in transit.

Furthermore, robust key management strategies, multi-factor authentication (MFA), and role based access control (RBAC) ensured that only authorized personnel could access the information, minimizing the risk of data breaches. To maintain resilience against evolving cyber threats, regular audits, penetration testing, and key rotations were conducted, ensuring that security measures remained up to date and effective. In securing the Integrated Crime Risk Assessment System (ICRAS), a comprehensive data sensitivity analysis was undertaken to assess potential risks and

vulnerabilities. Compliance with international data protection standards, including the General Data Protection Regulation (GDPR), further strengthened the legal and ethical foundation of the study. Additionally, the implementation of incident response plans and ongoing user training enhanced security awareness, equipping stakeholders with the knowledge to mitigate potential threats. Continuous monitoring, documentation, and accountability measures ensured that all data protection policies were not only enforced but also dynamically improved in response to emerging challenges. Through this multi-layered approach, the study upheld the highest standards of data privacy, security, and regulatory compliance, ensuring that participant confidentiality remained uncompromised while enabling the research to achieve its objectives with integrity and reliability.

3.4 Data Analysis

Data analysis is central to transforming the raw data collected in a research study into meaningful insights. For this study on crime in the Shah Alam district from 2015 to 2021, the data analysis methods comprise a blend of quantitative and qualitative techniques. These methods offer a holistic approach, enabling the study to reveal statistical crime trends and, at the same time, capture the perceptions, experiences, and explanations of crime dynamics from various stakeholders. This chapter will outline the statistical and thematic analysis methods employed in the study, addressing both the methodological and practical aspects of data interpretation. Statistical analysis in crime research allows for the discovery of patterns, trends, and relationships that would otherwise remain hidden in large datasets. By employing robust statistical techniques, this research aims to provide evidence-based insights that can inform policy and practice in crime prevention and law enforcement in Shah Alam.

i. Data Preparation and Cleaning

The first step in statistical analysis is preparing and cleaning the data. In crime analysis, the raw data is typically obtained from police reports, criminal records, and public safety databases. These records often contain inconsistencies, such as missing data, typographical errors, and redundant entries, which must be addressed before

analysis begins. In this study, crime incident data from 2015 to 2021 is cleaned using several steps:

a. Handling Missing Data: Missing data is a common issue in crime databases. Depending on the nature of the missing data, different approaches are used. For example, if a crime type or location is missing from a record, that entry may be excluded from certain analyses. Alternatively, statistical imputation methods, such as mean substitution or regression based imputation, may be used if the missing data is random and the dataset is large enough.
b. Standardization of Variables: Crime data often includes different formats for date, time, and location. To ensure consistency across the dataset, date and time formats are standardized, and categorical variables, such as crime types, are coded numerically for easier analysis.
c. Outlier Detection: Outliers can distort statistical results, particularly in crime data where a few extreme incidents (e.g., mass crimes or unusual spikes in crime activity) may skew the findings. Outliers are identified using methods like Z-scores or box plots and are treated carefully, either being included with justification or excluded to prevent misinterpretation.

ii. Descriptive Statistics

Descriptive statistics provide a general overview of the dataset, summarizing the main characteristics and trends in the crime incidents from 2015 to 2021.

Key metrics used include:

a) Frequency and Percentages: The total number of crime incidents within each category (property crimes, violent crimes, etc.) is calculated, along with their respective percentages. This step helps identify which types of crimes are more prevalent in Shah Alam and how their rates change over the 7-year period.
b) Central Tendency and Dispersion: Measures like the mean, median and mode are used to summarize the typical crime rate in the district, while measures of dispersion (standard deviation and variance) help understand the variability in

crime occurrences. For example, the mean number of crimes per year provides an idea of general crime frequency, while the standard deviation shows whether crime rates fluctuate significantly across years [.

c) Cross tabulation: Cross-tabulation involves comparing two categorical variables. For instance, crime types (e.g., theft, assault) are cross tabulated with locations (e.g., commercial areas, residential zones), providing insight into how different crimes are distributed across Shah Alam's diverse geographical landscape.

iii. Inferential Statistics

While descriptive statistics offer a broad overview, inferential statistics enable the study to make broader generalizations about the population based on the sample data. This section includes the following methods:

a. ANOVA (Analysis of Variance): ANOVA is employed to compare crime rates across different years or geographic regions. This technique helps in identifying whether differences in crime rates between neighborhoods or across different years are statistically significant [210].
b. Regression Analysis: Multiple regression analysis is utilized to explore the relationships between crime rates and several independent variables, such as population density, average income, and educational attainment. By analysing these predictors, regression models help identify which factors most strongly influence crime trends, offering evidence-based policy recommendations [211].

3.4.1 Overview of Thematic Analysis

While statistical analysis focuses on the numerical trends in the data, thematic analysis offers deeper insights into the qualitative aspects of crime in Shah Alam. This method is particularly valuable for interpreting the narratives and perceptions of individuals involved in or affected by crime, such as law enforcement personnel, community leaders, and victims. The thematic analysis process for this research comprises six interconnected stages which are data collection and familiarization, initial

coding, theme development, reviewing and refining themes, defining and naming themes, and interpretation and integration.

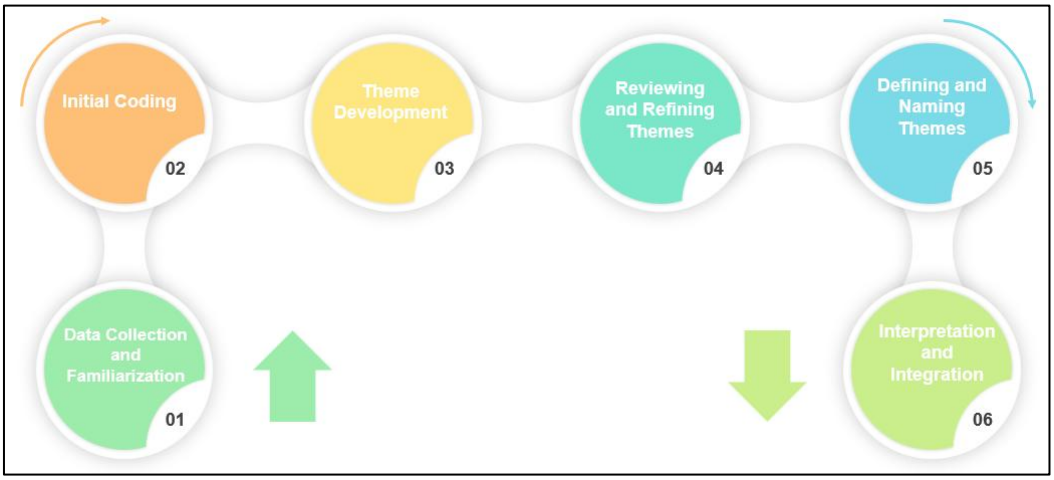


Figure 3.16 Thematic Analysis Flowcharts

The process begins with data collection and familiarization, a critical phase where researchers immerse themselves in the raw data to gain a deep understanding of its content and context. This step often involves meticulous reading and re reading of transcripts, interview notes, or survey responses, facilitating the identification of recurring ideas and notable details [212]. Familiarization is not merely an administrative task but a foundational exercise that enables researchers to connect with the data on a meaningful level, laying the groundwork for subsequent analytical rigor. Following this, the initial coding phase is undertaken, where the data is systematically fragmented into smaller, meaningful units. These codes serve as labels that encapsulate the essence of specific data segments, allowing researchers to organize and categorize information effectively. This step demands a balance between open mindedness and critical analysis, as researchers must remain attuned to both expected and emergent insights. The codes generated in this phase act as building blocks for the more complex process of theme development, where related codes are grouped to form broader patterns that reflect underlying meanings within the data. Thematic development requires an iterative mind set, where initial ideas are constantly evaluated, refined, and expanded to capture the depth and breadth of the data accurately. The process advances to reviewing and refining themes, a stage characterized by rigorous scrutiny to ensure that the themes identified are coherent, distinct, and substantively supported by the data. This phase is crucial for enhancing the credibility of the analysis, as themes must resonate with the

data's overall narrative while avoiding overlap or redundancy. Researchers often revisit the data to validate their themes, confirming that they reflect the participants' voices and the study's objectives. Once the themes are solidified, they are defined and named, a step that involves crafting precise and evocative descriptions that encapsulate the essence of each theme. This naming process is more than a technical exercise; it is an intellectual endeavour that demands creativity and analytical precision to communicate the themes' significance effectively. Finally, the process culminates in the interpretation and integration phase, where the identified themes are synthesized into a cohesive narrative that aligns with the study's research questions and theoretical framework. This step transforms raw data into actionable insights, providing a comprehensive understanding of the phenomena under investigation. The integration of themes with existing literature ensures that the analysis contributes to the broader academic discourse, offering both theoretical and practical implications. Throughout this process, thematic analysis is guided by transparency, reflexivity, and methodological rigor, ensuring that the findings are credible and ethically grounded [213]. This approach underscores the utility of thematic analysis in producing meaningful and actionable insights, particularly in complex and context-sensitive research areas. By adhering to established methodological frameworks and integrating perspectives from existing studies, this research demonstrates the transformative potential of thematic analysis in bridging qualitative depth with analytical precision.

3.4.2 Integration of Statistical and Thematic Analysis

By combining both statistical and thematic analyses, this research provides a multidimensional perspective on crime in Shah Alam, enhancing the overall validity and richness of the findings.

i. Methodological Triangulation: Methodological triangulation involves using multiple methods to study the same phenomenon, thereby strengthening the reliability of the findings. In this study, triangulation is achieved by comparing the results from statistical analyses with the narratives from thematic analysis. For instance, if regression analysis shows that unemployment is a significant

predictor of crime, thematic analysis can provide deeper context, explaining how economic stress leads to crime from the perspective of local residents.

ii. Complementary Insights: The integration of statistical and thematic analyses also allows for complementary insights. While statistical analysis may show a significant increase in property crimes in a specific year, thematic analysis can reveal public perceptions of that trend, such as increased fear of burglary or dissatisfaction with police response times. Together, these methods offer a fuller picture of the crime dynamics in Shah Alam.

iii. Contextualizing Findings: By integrating both quantitative and qualitative data, this study can better contextualize its findings. For instance, geospatial analysis might identify certain areas as crime hotspots, while thematic analysis reveals that these areas are perceived by residents as neglected or underserved by law enforcement. This integration helps explain not only where crime happens, but also why it happens in specific locations, offering more nuanced insights for policymakers.

3.4.3 Predictive Modeling Using Random Forest Algorithm

This study employed the Random Forest (RF) algorithm as the primary predictive modelling technique to model and evaluate crime risk patterns within the Integrated Crime Risk Assessment System (ICRAS). Random Forest was selected due to its strong capability in handling heterogeneous, non-linear, and moderately sized datasets, as well as its robustness to noisy and incomplete data, which are common characteristics of real-world crime datasets.

Crime data obtained from law enforcement agencies typically contain inconsistencies, missing entries, and variations in reporting practices. These characteristics introduce noise and uncertainty into the dataset. Ensemble learning techniques such as Random Forest are well-suited to such environments because they construct multiple decision trees using different subsets of the data and aggregate their predictions via majority voting. This process reduces model variance and minimizes the risk of overfitting that commonly occurs in single decision tree models. As a result, Random Forest produces more stable and generalizable predictions when compared to individual tree-based approaches.

The dataset used in this study consisted of 5,315 crime records collected between 2015 and 2021, representing violent and property crimes within the Selangor region. While this dataset provides sufficient information for machine learning analysis, it is considered moderate in size for predictive modelling purposes. Deep Learning approaches generally require substantially larger datasets to achieve optimal performance and stable convergence during training. When applied to moderate-sized datasets, deep learning models may experience overfitting and unstable generalization performance. In contrast, Random Forest has been shown to perform reliably on datasets of limited to moderate size while maintaining strong predictive accuracy and robustness.

Another important consideration in selecting the predictive modelling technique was model interpretability. The ICRAS system is designed to support Investigation Officers (IOs) in operational decision-making. Therefore, it is essential that the prediction outputs can be explained and interpreted in a meaningful manner. Random Forest provides built-in feature importance measures, enabling the identification of key predictors influencing crime risk levels. This transparency allows law enforcement officers to understand the reasoning behind the model's predictions and enhances trust in the system. Deep Learning models, on the other hand, typically operate as black-box models, making it difficult to interpret the relationships between predictors and outcomes without additional complex explanation techniques.

Computational efficiency and system deployment requirements also influenced the model selection. The ICRAS platform is intended for operational use within law enforcement environments where computational resources may be limited. Deep Learning algorithms generally require high-performance hardware such as Graphics Processing Units (GPUs) and longer training times. Random Forest models can be trained efficiently using standard computing infrastructure while still maintaining reliable predictive performance. This makes Random Forest more suitable for practical deployment within the ICRAS mobile-based framework. Table 3.3 presents a comparative evaluation of Random Forest and Deep Learning approaches with respect to criteria relevant to the ICRAS system.

Table 3.3
Comparison of Random Forest and Deep Learning for Crime Risk Prediction

Criteria	Random Forest	Deep Learning
Dataset size requirement	Suitable for small to moderate datasets	Requires large-scale datasets
Robustness to noisy data	High tolerance to noise and missing values	Sensitive to noisy and incomplete data
Model interpretability	High (feature importance available)	Low (black-box model)
Risk of overfitting	Low due to ensemble averaging	Higher with limited data
Training complexity	Moderate	High
Suitability for law enforcement deployment	High	Moderate to Low

Based on these considerations, Random Forest was determined to be the most appropriate predictive modelling technique for the ICRAS framework. In developing the predictive component of the system, the dataset was first cleaned and structured through the ETL process to ensure data integrity and consistency. Key features such as population density, unemployment rate, police response time, and historical crime incidents were selected as predictor variables. These variables represent socio-economic, operational, and spatial factors known to influence crime patterns. The Random Forest model was trained to classify and estimate crime risk levels by learning the relationships between these predictor variables. Each decision tree within the forest generated an individual classification result, and the final model prediction was obtained through majority voting across all trees. This ensemble approach enabled the model to capture complex interactions among socio-economic and environmental variables that influence crime risk patterns. Feature importance analysis generated by the Random Forest model was examined to determine the relative contribution of each predictor variable. This analysis enhanced the interpretability of the model and provided practical insights for law enforcement decision-making. Variables that consistently appeared as strong predictors were used to guide resource allocation and prioritization of preventive measures in high-risk areas. Model validation was carried out using standard performance metrics, including accuracy, precision, recall, and the Area Under the Receiver Operating Characteristic Curve (ROC-AUC). These metrics provided a comprehensive evaluation of the model's predictive capability and its effectiveness in distinguishing between different levels of crime risk. The evaluation results indicated

that the Random Forest algorithm produced stable and reliable predictions, establishing a strong analytical foundation for the ICRAS system.

The integration of Random Forest into the ICRAS framework enhanced the analytical capability of the platform by enabling dynamic interpretation of crime data collected from multiple sources. The model supported the generation of real-time risk assessments and evidence-based recommendations for Investigation Officers. Quantitative prediction outputs were interpreted alongside qualitative findings obtained from user feedback, providing a comprehensive understanding of crime risk patterns. The combination of ETL-based data integration, Random Forest predictive modelling, and user-oriented interpretation contributes to the development of a reliable and operationally practical system for proactive crime risk assessment within the Royal Malaysian Police environment.

3.4.4 System Challenges

Table 3.4
Challenges and Solutions in the Development of ICRAS

Challenge	Description	Solution
Data Standardization	Integrating crime data from various sources (police records, public reports, etc.) with different formats.	Utilized ETL (Extract, Transform, Load) processes to convert disparate data into a consistent format.
Real-Time Data Processing	Need for real-time processing of large volumes of crime data.	Developed advanced data pipelines optimized for high throughput and low latency to handle streaming data.
Predictive Accuracy	Ensuring accuracy of predictive models by integrating historical crime data with real-time environmental factors.	Employed machine learning techniques (e.g., random forests, neural networks) to continuously update risk assessment models.

3.5 Entity Relationship Diagram (ERD) and Data Model

An Entity Relationship Diagram (ERD) serves as a critical tool in representing the structure of a database by illustrating the entities within the system, their attributes, and the relationships between them. It provides a visual framework for understanding data flow and interconnections, offering both a conceptual and logical view of the database architecture [239]. The ERD typically incorporates entities such as "User,"

"Product," or "Transaction," and links them through relationships such as "one-to-many" or "many-to-many," which signify how the entities interact with one another [214]. A well-designed ERD is integral to ensuring that a database system is scalable, efficient, and easily maintainable, as it provides clarity for developers and database administrators. The Data Model is a broader representation that includes not only the structure but also the business rules and constraints that govern data usage and interactions. A data model can take several forms, ranging from high-level conceptual models that abstract away technical details to detailed logical and physical models that provide more granular insights into implementation specifics [215]. Data models typically incorporate the ERD but extend it to include more comprehensive elements such as data flow diagrams, normalization processes, and schema definitions. In the context of crime risk assessment systems like the Integrated Crime Risk Assessment System (ICRAS), the data model becomes instrumental in defining how various crime-related data ranging from incident reports, geographic locations, demographics, and environmental factors are structured and interconnected. The entity relationship framework provides clarity on how these data points are related, facilitating an understanding of complex crime patterns and aiding in predictive analytics [216]. For example, in an ICRAS implementation, entities such as "Crime Incident," "Suspect," and "Location" would be interconnected through relationships, allowing for sophisticated analysis and risk prediction. This data modelling approach is crucial in enhancing the accuracy and reliability of predictive outcomes. Furthermore, the use of normalization in data modelling ensures that redundancy is minimized, and data integrity is maintained, making the model both efficient and accurate. For instance, the normalization process might result in separating related data into multiple tables (e.g., separating crime data from demographic data), which can then be linked back together using foreign keys. This is a standard practice in developing crime assessment databases, which helps manage large-scale datasets and complex relationships efficiently. As ICRAS is designed to integrate data from multiple sources, the ERD and data model collectively support its ability to assess crime risks in a dynamic and real time context. A robust data model not only streamlines the process of data integration but also enhances the system's ability to perform advanced analytics, such as predictive modelling, by ensuring that the underlying data structure is aligned with the system’s analytical goals.

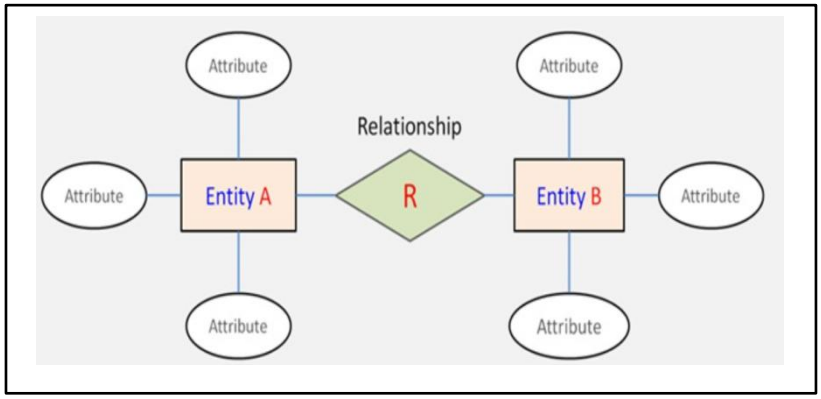


Figure 3.17 Entity Relationship Diagram (ERD)

3.5.1 Entities and Attributes

Table 3.5
Entities and Relationships in ICRAS

Entity Name	Attributes	Relationship
CrimeIncident	CrimeID (PK), IncidentDate, IncidentTime, CrimeType, Description, LocationID (FK), OfficerID (FK), SeverityLevel, FinancialLoss	One-to-Many with Location, One-to-Many with Officer
Location	LocationID (PK), Zone, Address, Latitude, Longitude	One-to-Many with CrimeIncident
Officer	OfficerID (PK), Name, BadgeNumber, Rank, DepartmentID (FK)	One-to-Many with CrimeIncident, Many-to-One with Department
Department	DepartmentID (PK), DepartmentName, LocationID (FK)	One-to-Many with Officer, One-to-One with Location
Victim	VictimID (PK), Name, Age, Gender, Address, ContactNumber	Many-to-Many with CrimeIncident (through CrimeVictim)
Crime Victim	CrimeID (PK, FK), VictimID (PK, FK)	Many-to-Many between CrimeIncident and Victim
Suspect	SuspectID (PK), Name, Age, Gender, Address, CriminalRecord	Many-to-Many with CrimeIncident (through CrimeSuspect)
Crime Suspect	CrimeID (PK, FK), SuspectID (PK, FK)	Many-to-Many between CrimeIncident and Suspect
Witness	WitnessID (PK), Name, Age, Gender, Address, ContactNumber	Many-to-Many with CrimeIncident (through CrimeWitness)
Crime Witness	CrimeID (PK, FK), WitnessID (PK, FK)	Many-to-Many between CrimeIncident and Witness
Evidence	EvidenceID (PK), CrimeID (FK), EvidenceType, Description, DateCollected	One-to-Many with CrimeIncident
Analysis	AnalysisID (PK), CrimeID (FK), AnalysisDate, AnalystID (FK), Findings	One-to-Many with CrimeIncident, One-to-Many with Analyst

Entity Name	Attributes	Relationship
Analyst	AnalystID (PK), Name, Specialty, DepartmentID (FK)	One-to-Many with Analysis
Crime Prevention	PreventionID (PK), CrimeType, Strategy, Description, Effectiveness	Linked conceptually to CrimeType for strategy recommendations

3.5.2 Entity Relationship Structure and Analytical Framework

The Integrated Crime Risk Assessment System (ICRAS)is underpinned by a robust analytical framework, grounded in well-defined entities and relationships. At its core is the CrimeIncident entity, which functions as the focal point for recording various incidents, with attributes such as the unique CrimeID (serving as the primary key), the date and time of the incident, and critical details like the crime type, description, and location. Each crime incident is linked to its location through a one-to-many relationship, whereby multiple crime incidents may occur at a single location. This relationship is established using the Location entity, which contains attributes such as LocationID (primary key), geographical details (latitude and longitude), and the zone in which the location falls. The Officer entity plays a pivotal role, recording details of the officer managing the incident, including the OfficerID (primary key), badge number, rank, and the department to which the officer belongs. Each officer is linked to a crime incident through a one-to-many relationship, indicating that one officer can be responsible for multiple crime incidents. Additionally, the officer's relationship with the Department entity is characterized by a many-to-one relationship, meaning that multiple officers belong to one department. The Department itself is identified by DepartmentID (primary key) and is associated with a specific location through a one-to-one relationship, allowing each department to manage incidents occurring in a particular geographical area.In terms of participants in the crime, the framework includes the Victim, Suspect, and Witness entities, each connected to the crime incident via intermediary tables. The CrimeVictim, CrimeSuspect, and CrimeWitness entities facilitate many-to-many relationships between the CrimeIncident and the associated people (victims, suspects, and witnesses). Each intermediary entity links the respective participants using composite keys, ensuring that multiple victims, suspects, or witnesses can be linked to a single crime incident, while each person can be associated with multiple crime incidents as well. The Evidence entity is another critical component, capturing physical or digital materials related to the incident. It maintains a one-to-many

relationship with CrimeIncident, allowing for the collection of multiple pieces of evidence for a single incident. The entity stores attributes such as EvidenceID (primary key), the type of evidence, and the date it was collected. This evidence is then analyzed through the Analysis entity, where each analysis is tied to a specific crime incident and performed by an analyst. A one-to-many relationship exists between Analysis and both CrimeIncident and Analyst, the latter storing details about the analysts, including their specialty and departmental affiliation.

Finally, the CrimePrevention entity, although not directly tied to crime incidents through relationships, plays a strategic role in crime risk assessment. It includes strategies based on crime type, with attributes capturing the effectiveness and description of each prevention strategy. This entity conceptually links to the CrimeIncident's CrimeType attribute, providing recommendations for future crime prevention efforts.

3.6 Programming Languages and Frameworks

The development of ICRAS involved multiple programming languages and frameworks that collectively ensured the system's robustness, responsiveness, and flexibility. The chosen technologies align with current web development trends, emphasizing rapid development and cross-platform compatibility.

i. PHP 8.2

Latest version of the popular server-side scripting language serves as the backend of ICRAS. PHP has long been known for its flexibility, ease of use, and vast ecosystem of libraries. PHP 8.2 introduced several significant features that enhanced the development of ICRAS, such as JIT (Just-In-Time) compilation and improved type safety, making the system more efficient and secure.

The JIT compilation mechanism allows PHP to execute certain types of code faster by converting PHP byte code into machine code. This feature significantly improved the performance of ICRAS, especially during large scale data processing tasks, such as analysing crime data or generating predictive reports.

Additionally, PHP 8.2's stricter type safety ensures that the system is more robust and less prone to errors. By enforcing strict typing, the development team

minimized issues related to type juggling, which can lead to bugs and security vulnerabilities in dynamic languages. This feature ensured data consistency, especially when handling complex datasets from multiple sources like crime reports, environmental data, and geospatial information.

PHP also supports extensive database integration features, which were used to connect the backend to MariaDB and Swagger API. [217] emphasized the importance of PHP in creating secure, dynamic web applications, especially when integrated with modern databases.

```php
<?php

declare(strict_types=1);

require 'vendor/autoload.php';
use Firebase\JWT\JWT;
use Firebase\JWT\Key;

class ICRASBackend {
    private PDO $db;
    private string $jwt_secret = "your_secret_key";

    public function __construct() {
        $dsn = "mysql:host=localhost;dbname=icras;charset=utf8mb4";
        $username = "root";
        $password = "";
        $options = [
            PDO::ATTR_ERRMODE => PDO::ERRMODE_EXCEPTION,
            PDO::ATTR_DEFAULT_FETCH_MODE => PDO::FETCH_ASSOC,
        ];
        $this->db = new PDO($dsn, $username, $password, $options);
    }
```

Figure 3.18 Secure Backend Implementation Using PHP 8.2

ii. Bootstrap 5

For the frontend development, Bootstrap 5 was used to ensure a responsive, clean, and intuitive interface for ICRAS users. Bootstrap is a popular open-source framework for building responsive, mobile-first websites. In ICRAS, Bootstrap 5 allowed the system to offer a consistent user experience across various devices, from desktop computers in police headquarters to mobile devices used by officers in the field. One of Bootstrap 5's key features are its flexible grid system, which enabled ICRAS to present data visualizations, charts, and tables in a structured and visually appealing manner. This grid system dynamically adjusts the layout based on the user's screen size, ensuring that users can access critical information whether they are working from a desktop monitor or a smartphone. Moreover, the pre built components in Bootstrap 5, such as navigation bars, buttons, modals, and form elements, sped up the development

process, allowing the team to focus on implementing core functionalities rather than spending time designing UI elements from scratch. Bootstrap 5 also removed jQuery as a dependency, making the application lighter and faster to load.

iii. JQuery

While Bootstrap 5 reduced its reliance on JQuery, this JavaScript library remained an integral part of ICRAS's interactive features. JQuery simplifies DOM (Document Object Model) manipulation, event handling, and asynchronous JavaScript requests, enabling dynamic behaviour in the system.

JQuery's simplicity and cross-browser compatibility made it an excellent choice for adding interactive elements to ICRAS, such as filtering data, real-time search capabilities, and updating sections of the webpage without requiring a full reload. This made the user experience smoother, particularly for officers who needed quick, real-time access to crime data without delays. [218] praised JQuery for its versatility in enhancing the functionality of modern web applications.

iv. JavaScript

In addition to JQuery, JavaScript played a key role in the development of ICRAS. While JQuery efficiently manages simpler DOM manipulations, JavaScript was used for more complex interactivity and custom logic. Specifically, JavaScript was employed to handle dynamic crime data interactions and user processes. This includes managing client-side form validation, ensuring that data entered ICRAS's interface is verified before being sent to the server. This approach reduces server load and improves the system's efficiency, particularly when handling large datasets from multiple users simultaneously. JavaScript also handles client-side form validation, improving the user experience by ensuring that data entered ICRAS's interface is checked before being sent to the server. This reduces server load and improves the system's efficiency, especially when dealing with large datasets from multiple users simultaneously.

v. AJAX (Asynchronous JavaScript and XML)

AJAX was extensively used in ICRAS to facilitate asynchronous communication between the client and server, enhancing the system's interactivity and responsiveness. By using AJAX, ICRAS can send and retrieve data from the server without reloading the entire page, enabling smooth updates and dynamic content loading. For example, when users request data from the system, AJAX allows for the retrieval of relevant information without interrupting the user experience. This approach ensures seamless interaction with the system, allowing for continuous data processing and analysis, which is essential for tasks requiring efficiency and minimal delays.

vi. MobileDetect (Responsive Design)

Given that ICRAS is used by law enforcement personnel both in the field and in the office, the system was designed to be fully responsive. MobileDetect, a lightweight PHP class for detecting mobile devices, was used to optimize the system's display and functionality based on the user's device.

By incorporating MobileDetect, ICRAS dynamically adjusts its interface to provide an optimal user experience for mobile devices, including smartphones and tablets. Whether officers are using ICRAS on a desktop workstation or a mobile phone, the system delivers a consistent, usable interface. Responsive design is critical in situations where officers need to access ICRAS on-the-go, especially during patrols or emergency responses.

3.6.1 Database: MariaDB 10

For the database management of ICRAS, MariaDB 10 was selected due to its robust performance, scalability, and compatibility with existing SQL databases. MariaDB, a popular open-source relational database, is highly optimized for performance while maintaining compatibility with MySQL, making it an ideal choice for managing large datasets.

Table 3.6
Relational Data Management in ICRAS using MariaDB

Aspect	Detail
Database Structure	MariaDB's relational database supports ICRAS's data relationships, including crime records, geographic data, environmental factors, and predictive analytics.
Entity Relationship Diagram (ERD)	The ERD for ICRAS consists of 14 entities and 64 attributes, mapping out key data points like crime reports, officer details, locations, suspect/victim info, and predictive analytics results.
Data Relationships	ICRAS performs complex queries by cross-referencing multiple entities, enabling meaningful crime risk assessments, such as generating crime statistics by geographic region over time.
Optimization and Indexing	ICRAS uses indexing on frequently searched fields (e.g., crime location, time, type), improving query response times. MariaDB 10's partitioning feature divides large tables, enhancing performance, especially for historical data.
Security Features	MariaDB includes role-based access control (RBAC), encryption, and auditing to protect sensitive crime data. Encryption ensures secure storage and transmission of personal details (e.g., suspects, victims).

3.6.2 Swagger API for Integration and Documentation

To facilitate communication between ICRAS and external systems, Swagger API was used for designing and documenting the RESTful APIs that enable external applications to interact with ICRAS. Swagger allows the system to have well-structured API documentation, making it easier for external developers and system administrators to integrate other services, such as geospatial mapping tools, emergency response systems, and data-sharing platforms from other agencies.

Swagger's auto-generated API documentation also ensures that as the ICRAS backend evolves, the API remains consistent and up to date, providing a standardized way for other systems to connect and exchange data securely.

3.7 Data Flow and Process in ICRAS

The data flow within ICRAS is designed to support real-time crime analysis and risk assessment. Data from multiple sources including crime reports, environmental factors, and public surveillance are continuously ingested, processed, and analysed to provide actionable insights. The system's architecture is built to handle both batch and real-time data, ensuring that predictions and assessments are always based on the most current information available.

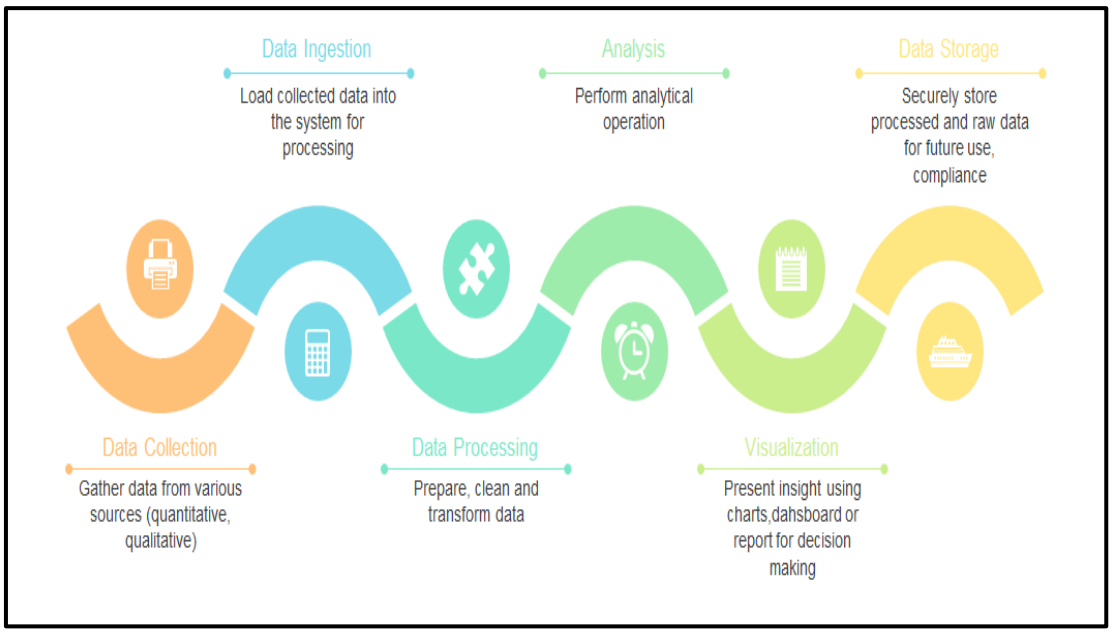


Figure 3.19 ICRAS Process

The data flow process is a foundational framework in research and analytics, guiding the transformation of raw data into meaningful insights. This process begins with data collection, where information is gathered from diverse sources such as sensors, databases, devices, and manual input. Modern advancements, such as IoT systems and automated collection mechanisms, have enhanced the efficiency and accuracy of this phase. Once collected, the data undergoes data ingestion, a process of integrating information into a centralized system for processing. This step leverages frameworks like Apache Kafka or Apache NiFi to enable seamless data transfer, whether in real time streaming or batch uploads, as discussed in research by [219]. Efficient ingestion also facilitates metadata tagging, adding context to the data for subsequent stages. The data processing phase transforms raw data into usable formats through cleaning and transformation. This step addresses inconsistencies such as missing values, duplicates, and outliers while ensuring compatibility with analytical models through techniques like standardization and normalization. [220] highlight the growing use of machine learning algorithms in automated data cleaning, reducing human intervention and improving accuracy. Additionally, ETL (Extract, Transform, Load) frameworks have been instrumental in streamlining this phase, as detailed by [221]. Following processing, the data enters the analysis stage, where statistical tools, machine learning models, and predictive algorithms uncover patterns and insights. This stage can involve clustering, regression, or time-series analysis, depending on the research objectives. Studies by [222] show that hybrid approaches, which combine

traditional statistical methods with AI-driven models, are becoming increasingly popular in domains like healthcare and finance. The integration of deep learning techniques has further expanded the possibilities for extracting actionable insights from complex datasets. The insights generated during analysis are then presented through visualization, which translates complex data into comprehensible formats such as charts, graphs, and interactive dashboards. Visualization tools like Tableau and Power BI enable users to interact with data dynamically, refining inputs and adjusting visual parameters. Research by [223] highlights the critical role of visualization in fostering collaboration between technical teams and decision-makers, ensuring that insights are actionable and widely understood. Finally, the data storage phase ensures the secure and systematic archiving of both raw and processed data. With the rise of cloud-based storage solutions like AWS and Azure, alongside distributed databases like Hadoop and MongoDB, this phase has evolved to prioritize scalability, redundancy, and compliance with regulations such as GDPR and HIPAA [224]. Techniques like data deduplication and compression further optimize storage efficiency, as noted by [225].

This six-step data flow process spanning collection, ingestion, processing, analysis, visualization, and storage provides a robust methodology for managing data in modern research. By integrating principles from diverse studies, this framework ensures the efficient transformation of raw data into insights, supporting decision-making and future innovations. Future research should focus on automating feedback loops and leveraging emerging technologies like edge computing to further refine this process and enhance its scalability and adaptability.

3.7.1 Security and Data Privacy Considerations

Given the sensitive nature of the data handled by ICRAS, data security and privacy were paramount concerns during development and deployment. The system incorporates multiple layers of security protocols to ensure the confidentiality and integrity of both historical and real-time data.



Figure 3.20 Security and Data Privacy Elements

i. Encryption: All data transmitted through ICRAS, whether between internal modules or between the system and external databases, is encrypted using Advanced Encryption Standard (AES-256). This ensures that even if data were intercepted during transmission, it would be unreadable without the correct encryption keys.

ii. Access Control: A role-based access control (RBAC) mechanism was implemented to restrict access to different parts of the system based on user roles. Officers, analysts, and administrative staff each have different levels of access, ensuring that sensitive data is only available to authorized personnel. The system also tracks all access attempts and modifications to data through an audit trail, providing accountability in case of data misuse or breaches.

iii. Anonymization: To comply with data privacy laws and protect the identities of crime victims and suspects, personal identifiers such as names, addresses, and contact details are anonymized in certain sections of the system. These anonymized records are used in higher-level analysis, such as predictive modelling, while the raw data remains accessible only to individuals with the proper clearance.

iv. Regular Security Audits: ICRAS undergoes periodic security audits conducted by independent cybersecurity firms. These audits test for vulnerabilities in the system's architecture and ensure that security protocols remain robust against emerging cyber threats. [226] underscore the importance of routine audits in maintaining the integrity of crime data systems, especially those operating in high risk environments

3.8 Pilot Study

3.8.1 Overview of the Pilot Study

Pilot studies play an essential role in system development and validation, particularly when implementing complex systems like the Integrated Crime Risk Assessment System (ICRAS). A pilot study acts as a preliminary exploration to verify the system's functionality, assess usability, and ensure the validity of data collection instruments before proceeding to full scale deployment . The implementation of ICRAS in the Shah Alam district requires robust validation to confirm that it meets the practical needs of law enforcement officers, crime analysts, and policymakers. This system was designed to predict and assess crime risks, allowing stakeholders to make data driven decisions. However, like any large scale technological deployment, it must be tested in a real-world scenario to identify any potential weaknesses or areas for improvement. In this regard, the pilot study provided the necessary framework for testing the system in a controlled environment, ensuring that it could handle real-time crime data, generate accurate risk assessments, and support users in their day-to-day tasks.

A successful pilot study reduces the likelihood of costly errors during full implementation and improves the likelihood of achieving the system's intended outcomes. It also serves as a crucial feedback loop for developers, allowing them to adjust the system based on actual user experiences. This process aligns with best practices in system development, particularly in risk sensitive domains such as crime prevention and law enforcement.

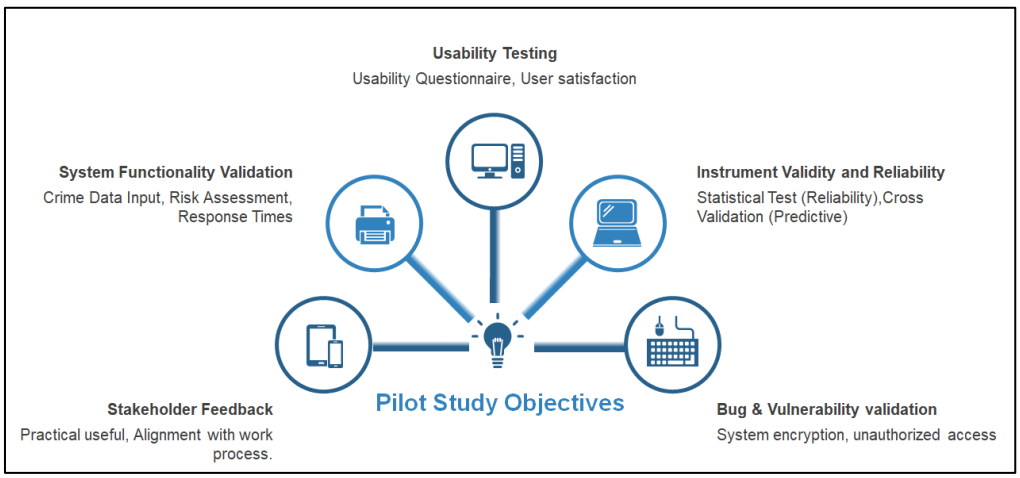


Figure 3.21 Overviews of Pilot Study Objectives for Evaluating ICRAS Performance and Usability

i. Validation of System Functionality: The study aimed to evaluate whether ICRAS performed as intended. This includes an assessment of key technical functions, such as crime data input, data analysis capabilities, and the generation of risk assessment reports. The system's ability to handle large volumes of data, particularly under stressful conditions, was tested to ensure scalability and reliability. Performance metrics, such as response times and error rates, were measured to provide a comprehensive understanding of the system's operational efficiency.

ii. Usability Testing: Usability is a critical factor in the success of any system, especially one designed for end-users with varying levels of technical expertise. The study assessed the ease of use of ICRAS through direct interaction with stakeholders. A usability questionnaire was administered to gauge user satisfaction and identify areas where the system's interface could be simplified or optimized.

iii. Instrument Reliability and Validity**:** Another core objective was to verify the accuracy of the instruments embedded in ICRAS, such as crime data collection tools and predictive models for crime risk assessment. Statistical tests were used to assess the reliability of these instruments, while cross-validation was applied to ensure the accuracy of the system's predictive capabilities.

Table 3.7
Roles and Responsibilities of Participants in the ICRAS Pilot Study Evaluation

Participant Group	**Role in ICRAS Evaluation**	**Key Responsibilities**
Law Enforcement Officers	Primary users of ICRAS, relying on it for crime risk assessments and decision-making.	- Assess ease of accessing crime data - Evaluate usability and efficiency for real-time resource allocation
Crime Analysts	Experts in crime data input and analysis, focusing on the accuracy of predictive models and report generation.	- Evaluate data analysis functions - Ensure reports are accurate and actionable
IT Personnel	Backend managers responsible for system performance, including database management, server maintenance, and security.	- Monitor data storage and processing speed - Identify and resolve technical issues, especially cyber security

Total of 30 participants were selected for the study, with equal representation from each stakeholder group. This ensured a balanced perspective on the system's functionality and usability across different user categories. The decision to use purposive sampling was guided by the need to include participants with a deep understanding of crime data systems and their operational contexts. According to [227], purposive sampling is appropriate in cases where specialized knowledge is required to evaluate complex systems.

Table 3.8
Summary of Research Methodology for Integrated Crime Risk Assessment System (ICRAS)

No.	Description	Justification
1. Research Design	Mixed Method approach combining quantitative and qualitative methodologies	Ensures a comprehensive understanding of crime dynamics from both statistical and human perspectives.
1.1 Quantitative Insights	Uses large datasets to identify crime trends, patterns, and temporal/spatial factors (e.g., frequency of crimes over time).	Provides objective, data-driven insights into crime distribution and changes over time.
1.2 Qualitative Insights	Interviews and focus groups to explore contextual and socio-cultural factors influencing crime (e.g., perceptions of law enforcement, community trust).	Captures the human experiences and socio-cultural factors behind crime patterns.
1.3 Ethical Considerations	Compliance with ethical standards (e.g., Personal Data Protection Act 2010).	Ensures the protection of participant rights and the integrity of data throughout the research process.
2. Mixed Methods Approach	Combines both methods to provide a richer, more comprehensive understanding of crime dynamics.	Leverages the strengths of both approaches to ensure a balanced and well-rounded analysis.
2.1 Quantitative Component	Identifies crime patterns, prevalence, and spatial distributions using statistical techniques (e.g., regression, correlation).	Provides robust, generalizable data that helps identify trends and correlations in crime patterns.
2.2 Qualitative Component	Offers deeper insights into the human experiences behind crime patterns (e.g., through interviews with victims and law enforcement officers).	Provides valuable context and helps explain the 'why' behind the trends observed in quantitative data.

No.	Description	Justification
3. ICRAS Development	Focuses on developing the Integrated Crime Risk Assessment System (ICRAS) using mixed-methods research.	Ensures that the system design incorporates both empirical data and community input, improving its real-world relevance.
3.1 Problem Definition	Identifies the need for a systematic crime risk assessment framework in Malaysia	Sets the foundation for developing a model that addresses a critical gap in crime prevention in Malaysia.
3.2 Data Collection	Collects crime data (from police departments, socio-economic indicators, etc.) for model construction.	Ensures that all relevant data sources are considered, leading to an accurate and comprehensive crime model.
3.3 Model Development	Applies machine learning algorithms to access crime risks	Provides a data driven crime risk assessment model that can guide law enforcement interventions.
3.4 Evaluation	Tests the system's accuracy using performance metrics (e.g., precision, recall).	Ensures that the developed system meets the accuracy and reliability standards necessary for operational deployment.
4. Sampling Strategy	Systematic sampling used to ensure representativeness, minimizing bias and ensuring efficiency.	Provides a balanced, representative sample that accurately reflects crime data patterns across different years and locations.
4.1 Systematic Sampling	Every k-th crime incident selected from the ordered list (based on crime type, year, and location)	Reduces sampling bias and ensures the sample is evenly distributed across the dataset for better analysis.
5. Data Collection and Analysis	Structured quantitative and qualitative data collection, cleaning, and processing.	Ensures the data is comprehensive, accurate, and ready for meaningful analysis.
5.1 Quantitative Data	Data from police records, national crime statistics, and other government sources.	Provides the core data needed to identify trends and generate actionable crime risk insights.
5.1.1 Data Processing using ETL Framework	Systematic extraction, cleaning, transformation, and loading of data from multiple sources into a centralized repository.	Ensures data accuracy, consistency, and readiness for analysis, forming a reliable foundation for predictive modelling in ICRAS.
5.2 Qualitative Data	Interviews with law enforcement, community leaders, victims, and residents.	Provides a human element, allowing for the exploration of perceptions and experiences related to crime and safety.
6. Data Management and Ethics	Ensures ethical data handling and compliance with data protection laws, including encryption and anonymization techniques.	Guarantees that the research adheres to legal and ethical standards, fostering trust and ensuring data security.
7. Data Analysis Techniques	Uses both statistical analysis (descriptive, inferential) and thematic analysis for qualitative data.	Combines rigorous statistical methods with contextual insights, offering a more complete view of the crime situation.

No.	Description	Justification
7.1 Quantitative Data Analysis	Descriptive statistics (e.g., mean, median, mode), inferential statistics (e.g., regression, hypothesis testing).	Helps identify patterns and relationships within the data that are critical for crime prevention.
7.2 Qualitative Data Analysis	Thematic analysis to identify recurring themes in interviews and focus groups.	Captures in-depth narratives and helps explain the underlying factors influencing crime trends.
8. Integration of Quantitative & Qualitative Insights	Ensures triangulation, complementarity, and contextualization, increasing the reliability and depth of findings.	Enhances the overall validity and depth of findings by combining both data sources in a complementary manner.
8.1 Triangulation	Cross-verifying quantitative data with qualitative insights for enhanced validity.	Strengthens the credibility of findings by comparing results from multiple data sources.
8.2 Complementary Insights	Combining numerical data trends with personal narratives (e.g., how community members perceive the effectiveness of crime prevention strategies)	Enriches data interpretation by including personal perspectives, ensuring findings reflect both statistical trends and human experiences.
9. Methodological Considerations	Addresses challenges such as data integration, resource management, and maintaining data quality	Ensures the research process remains valid, reliable, and adaptable by anticipating and mitigating methodological challenges, thereby improving the overall robustness and credibility of the ICRAS development.
9.1 Data Integration Challenges	Effectively combining qualitative and quantitative data, ensuring consistency	Ensures that both data types work together cohesively, avoiding contradictions or biases.
9.2 Resource Management	Managing time, personnel, and funding to ensure a smooth data collection and analysis process.	Maximizes efficiency and ensures the research stays within its allocated resources.
10. ICRAS System Development & Tools	Technical implementation using technologies like PHP 8.2, Bootstrap 5, and JQuery for backend and frontend development	Ensures that the system is built with modern tools, allowing for a responsive, secure, and scalable platform.
10.1 Backend (PHP 8.2)	PHP used for backend, with improvements like JIT compilation for performance.	Improves performance and scalability, making the system faster and more reliable.
10.2 Frontend (Bootstrap 5)	Ensures a responsive, mobile-first interface	Guarantees that users, including law enforcement officers, can access the system from any device, enhancing usability.
10.3 Interactive Features (JQuery)	Adds dynamic, real-time interactivity for filtering data and enhancing user experience.	Improves user experience by providing real-time data filtering and interaction, streamlining workflows for officers.

CHAPTER 4
RESULT AND DISCUSSION

4.1 Summary of Results

This chapter presents the findings from the research conducted to develop the Integrated Crime Risk Assessment System (ICRAS) for the Royal Malaysian Police (RMP), based on data collected through questionnaires, surveys, and interviews with RMP users, particularly Investigation Officers (IOs). The results from the questionnaire and survey revealed key insights into the technological preferences of RMP personnel, particularly the widespread use of Android devices among IOs, which guided the decision to develop the ICRAS app for Android. However, recognizing the potential for broader use, the app was designed to be hybrid, ensuring it could also be accessible to Apple users in the future. Interviews with IOs further highlighted challenges they face in their investigative work, such as limited access to real time crime data and a lack of predictive tools for risk assessment. These findings reinforced the necessity for ICRAS to provide a more integrated and efficient solution. To address these needs, key components of the system were developed, starting with the Risk Appetite Index, which classifies crime risks based on multiple factors, followed by the Risk Matrix, which visualizes the risks to help officers prioritize interventions. The development process also included the creation of the Entity Relationship Diagram (ERD), which mapped the relationships between various entities within the system to ensure smooth data flow and integration. Following the ERD process, the system then moved into the HTML development phase, where the structure and layout of the ICRAS web interface were designed to ensure ease of use and efficient data presentation. Subsequently, the development of the ICRAS mobile app began focused on providing a user friendly, real time crime risk assessment tool for IOs. The app, designed for Android devices, incorporates the functionalities required by RMP personnel, such as real time crime updates and the ability to generate crime reports quickly. Once the app was developed, it went through a rigorous validation process, including a pilot test involving a selected group of RMP officers (Appendix 3). The pilot test assessed the app's functionality, usability, and overall effectiveness in real world conditions. Feedback from the pilot test, along with performance metrics, was used to fine tune the app, ensuring it met the

practical needs of IOs. This chapter will tabulate all results from each step of the development process, including feedback from the pilot test and final validation, providing a comprehensive overview of how ICRAS was refined and validated to meet the operational needs of the RMP.

4.2 Data Collection Method

4.2.1 Questionnaire

A total of 30 questionnaires were administered to Royal Malaysian Police (RMP) users, specifically targeting Investigation Officers (IOs). The response rate was 96.67%, with 29 out of the 30 selected participants completing the questionnaire. This high response rate is indicative of the strong interest and engagement among the IOs regarding the development of a crime risk assessment system that could potentially enhance their work. The questionnaire was designed to gather insights into their preferences for mobile platforms (Android or Apple), as well as the specific features they would find most valuable in a crime risk assessment system. The findings from the questionnaire clearly reveal that the majority of IOs prefer using Android devices, with 90% of respondents indicating that they primarily use Android based smartphones for their work. This is significantly higher than the 10% who reported using Apple devices. The data collected also highlighted several key features that RMP personnel deem important for a crime risk assessment system. These include real time crime updates**,** crime risk visualization, and user interface satisfaction. The detailed breakdown of these preferences is presented in the table below:

Table 4.1
Preferences for Mobile Platform and Key Features in the Crime Risk Assessment System

Feature	Android Users (%)	Apple Users (%)
Real time Crime Updates	90	10
Crime Risk Visualization	85	15
User Interface Satisfaction	92	8

From the data, it is clear that Android users overwhelmingly favoured features such as real time crime updates (90%) and crime risk visualization (85%). These results

suggest that IOs working on Android devices would greatly benefit from a system that delivers timely and accessible crime data, along with tools that help them visualize crime trends and patterns. Furthermore, 92% of Android users expressed satisfaction with the user interface, indicating a preference for a simple, intuitive design that allows them to access and interpret crime data with ease. While the data shows a strong preference for Android devices, the questionnaire responses also highlighted the potential for Apple users to benefit from the system. Despite the smaller percentage of Apple users (10%), the hybrid design of the ICRAS app ensures that it can be accessed by both Android and Apple users, making it adaptable to a wider range of devices and extending its usability across the RMP. The results of the questionnaire provide valuable guidance for the development of the ICRAS system, confirming that focusing on an Android based application is the most effective approach given the current device usage trends within the RMP, while still leaving room for future expansion to support Apple users as well.

4.2.2 Survey

In addition to the questionnaire (Appendix 2), a survey was conducted to gather broader and more quantifiable insights into the operational needs of the Royal Malaysia Police (RMP), specifically in relation to crime risk assessment. This survey was intended to complement and validate the qualitative findings collected earlier by offering a wider perspective on current operational limitations and future system expectations. A total of 50 participants took part in the survey, including officers, system administrators, analysts, and personnel involved in operational and strategic planning across various units of the RMP. At present, the RMP does not utilize any dedicated crime risk analysis application or integrated assessment platform. The only existing digital infrastructure in place is the e-reporting system, which is primarily designed for the submission and management of public police reports. This system, while useful for administrative purposes, does not support crime trend analysis, predictive policing, or real time risk assessment. Consequently, much of the risk evaluation and operational planning still relies on manual interpretation, experience-based judgment, and fragmented information sources. The main objective of this survey was to assess how these gaps affect daily policing operations and to explore the perceived value of implementing an Integrated Crime Risk Assessment System

(ICRAS). The survey aimed to identify specific areas where technological intervention is urgently needed, such as data accessibility, mobile functionality, predictive analytics, and inter-agency integration. Respondents were asked to rate their level of agreement with several statements representing key operational challenges and desired features in a future crime risk system. The data collected from this instrument offers quantifiable evidence of system needs and end user expectations.

The summarized results in Table 4.2 reflect a high level of support for the development of ICRAS. Participants expressed strong interest in real time access to crime data, user friendly analytical tools, centralized integration, and mobile support for field officers, and predictive analytics for proactive policing. These findings serve as a critical reference point in shaping the functional design and justification for ICRAS within the RMP.

Table 4.2
Summary of Survey Findings on Crime Risk Assessment Needs and ICRAS Expectations

No	Survey Statement / Key Findings	Percentage (%)	Interpretation
1.	Absence of real time tools for crime risk assessment	98	Indicates a critical need for real-time data access in ICRAS for faster operational decision-making
2.	Strong interest in predictive analytics for crime forecasting	92	Encourages the inclusion of AI-driven features in ICRAS for proactive policing
3.	Lack of mobile access to crime intelligence during field operations	83	Supports the development of mobile-accessible features within ICRAS
4.	Belief that a centralized system like ICRAS would improve operational efficiency / strategic planning	90	Confirms strong stakeholder support for ICRAS implementation

These results reveal a clear and consistent pattern of operational shortcomings that constrain the Royal Malaysia Police's (RMP) current capacity to perform efficient and informed crime risk assessments. A substantial 98% of respondents reported the absence of real time tools, which highlights the critical need for ICRAS to provide live data access, timely information updates, and practical decision support features that can assist officers during investigative and operational tasks. Furthermore, 92% of respondents expressed strong interest in predictive analytics, not for early threat

detection, but to function as a decision support system that helps investigation officers better understand crime developments, manage cases more effectively, and plan investigative strategies based on available data. In relation to field operations, 83% of respondents indicated the lack of mobile access to crime intelligence, underlining the importance of ensuring ICRAS is mobile compatible so that officers can access relevant information directly from their devices while on duty. Additionally, 90% of participants believed that the implementation of a centralized and integrated system like ICRAS would enhance operational efficiency and strategic planning, by improving coordination, reducing data fragmentation, and enabling more consistent and informed decision making. Taken together, these findings strongly support the development of ICRAS as a comprehensive and responsive system tailored to the operational realities and investigative needs of the RMP.

4.2.3 Interview Data

To complement the survey and questionnaire findings, in depth interviews were conducted with 30 Investigation Officers (IOs) from the Criminal Investigation Department (CID) of the Royal Malaysia Police (RMP). As the primary agency responsible for criminal investigations, the CID plays a vital role in uncovering facts, analysing evidence, and solving cases across various types of crime. Investigation Officers serve as the frontline personnel in this process, and their insights are critical for ensuring that any technological system such as the proposed Integrated Crime Risk Assessment System (ICRAS) meets operational needs effectively. To gain a balanced view across experience levels, the respondents were divided equally into two categories: 15 junior IOs (with less than 5 years of service) and 15 senior IOs (with more than 5 years of service). The interviews aimed to explore the practical challenges faced during investigations, their views on existing systems, and the specific features they expect from ICRAS. The results revealed a series of shared and distinct concerns, summarized in Table 4.3 below.

Table 4.3
Summary of Interview Findings from Junior and Senior CID Investigation Officers on System Needs and Challenges

No	Theme	Junior IO's	Senior IO's	Relevance to ICRAS
1.	Need for improved data accessibility	13 (87%)	13 (87%)	Both junior and senior officers struggle with accessing case related information quickly.
2.	Absence of real time crime risk assessment tools	12 (80%)	13 (87%)	Real-time data and risk evaluation tools are seen as vital for prioritizing investigative work.
3.	Importance of user friendly systems	14 (93%)	13 (87%)	Real-time data and risk evaluation tools are seen as vital for prioritizing investigative work.
4.	Need for mobile accessibility during field investigations	12 (80%)	13 (87%)	Both groups want access to crime intelligence while on the move or at the crime scene.
5.	Frustration with fragmented or manual data systems	11 (73%)	13 (87%)	Both groups want access to crime intelligence while on the move or at the crime scene.
6.	Support for a centralized all in one investigative platform	14 (93%)	13 (87%)	Nearly all respondents support a single platform for managing case data, risk insights, and reporting.

Based on the interview data from 30 CID Investigation Officers, a number of consistent themes emerged that highlight shared operational challenges and critical system needs. The most strongly supported theme was the need for a centralized, all in one investigative platform, cited by 93% of both junior and senior officers. This reflects a widespread frustration with having to access multiple disconnected systems, and a shared belief that consolidation would significantly improve investigation efficiency and case management. Similarly, 87% of all respondents expressed the need for improved data accessibility, noting delays in retrieving case related information and emphasizing the importance of having quick access to accurate data to support timely decisions. The absence of real time crime risk assessment tools was highlighted by 83% overall slightly more among senior officers (87%) than juniors (80%) underscoring the value of lives information and risk indicators for prioritizing and managing investigations more effectively. The importance of user friendly systems was also emphasized, with 93% of junior officers and 87% of senior officers agreeing that current tools are often difficult to navigate. This shows a strong preference especially among less experienced officers for simplified interfaces that allow for fast learning and

minimal training. In addition, 80% of all officers indicated the need for mobile access to crime intelligence, pointing out that investigations often take place outside the station and require real time updates in the field. Finally, 73% of both groups expressed frustration with fragmented or manual systems, describing them as inefficient and error prone. These consistent responses across all ranks clearly reinforce the case for ICRAS to be developed as a real time, mobile compatible, user centric, and fully integrated digital platform to meet the operational realities of CID officers.

4.2.4 Results of the ETL Process

The application of the ETL framework was central in converting large volumes of raw and fragmented data into a clean, structured, and analysis-ready format. This step proved critical in overcoming the inherent challenges of working with data from multiple, inconsistent sources. Beyond merely standardizing information, the process ensured that every dataset used in the study met rigorous standards of accuracy, consistency, and analytical usability an essential requirement for the reliability of the Integrated Crime Risk Assessment System (ICRAS). A key achievement of the ETL implementation was the substantial improvement in data accuracy and consistency. During the transformation phase, systematic checks were conducted to detect and resolve common issues such as missing entries, duplicates, and misclassified records. Variations in formats particularly in date, time, and geographical attributes were corrected to create uniformity across datasets. This process of refinement was not simply a technical necessity but a vital step in ensuring that the analytical outcomes derived from the data could be trusted. Reliable and comparable data formed the backbone for subsequent statistical modeling and risk analysis, minimizing the possibility of error propagation during interpretation. The integration of external datasets represented another major contribution of the ETL process. By combining crime data with socio economic and demographic indicators, the analytical depth of the study was notably enhanced. These additional layers of context covering factors such as population density, income level, and educational attainment allowed for a richer exploration of crime dynamics within Shah Alam. Such integration provided valuable insight into the socio structural elements influencing crime patterns, improving the predictive capacity of ICRAS and enabling more context aware risk assessments.

Normalization also played a crucial role, particularly for temporal and spatial data. Crime records often differed in time granularity or geographic reference systems, which made direct comparison challenging. Through normalization, time intervals were standardized and spatial coordinates aligned under a unified reference framework. This adjustment was indispensable for generating consistent, comparable analyses across different datasets and for enabling spatiotemporal modelling key components in identifying crime trends and forecasting potential hotspots. Equally important was the improvement in data accessibility and real time functionality. Once the processed data was loaded into a centralized repository, it became significantly easier to query and retrieve. The database structure was optimized for performance, allowing RMP officers and analysts to access up-to-date information almost instantaneously. This capability not only reduced delays in analysis but also facilitated proactive and evidence-based decision making reflecting the core aim of ICRAS to support real time policing and operational efficiency. Overall, the ETL process markedly enhanced the quality, coherence, and analytical readiness of the data. Through systematic cleaning, integration, and normalization, the datasets were transformed into a dependable foundation for advanced modeling and visualization. The process ensured that errors and inconsistencies were minimized, while relevant contextual variables were maximized. This transformation was not only technical but strategic it laid the groundwork for accurate risk assessment, effective resource allocation, and data-driven policing.

Furthermore, the ETL framework proved instrumental in bridging raw data with actionable intelligence. By resolving long standing issues of fragmentation, inconsistency, and inaccessibility, it allowed ICRAS to operate on a strong empirical footing. The improved data environment enabled law enforcement officers to make more informed, timely, and targeted interventions marking an important step toward a more predictive and intelligence led policing approach in Malaysia.

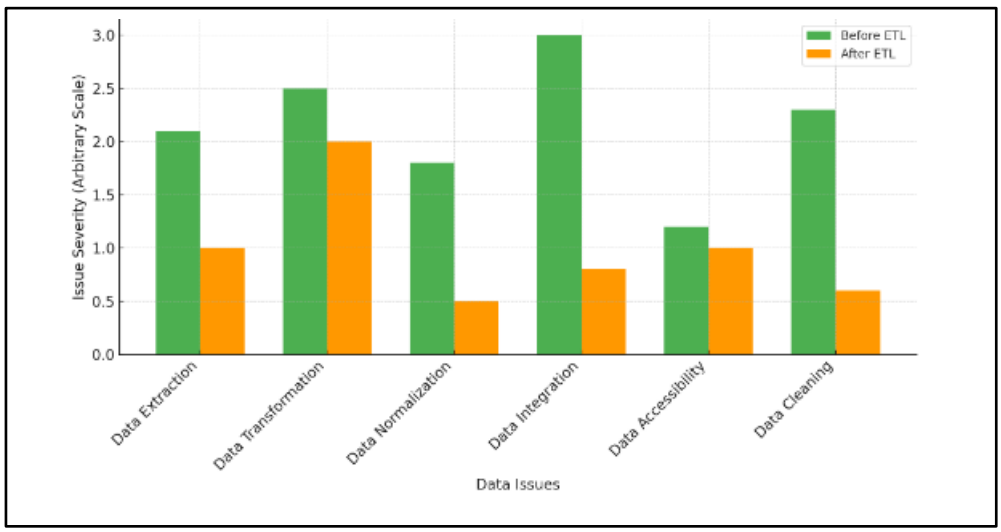


Figure 4.1 ETL Process Results Comparison

The updated bar chart illustrates a clearer comparison of data issue severity before and after the implementation of the ETL (Extract, Transform, Load) framework. In the chart, the green bars indicate the level of data issues prior to the ETL process, while the orange bars represent the condition after the process was applied. Overall, the ETL framework substantially reduced the severity of most data problems. In particular, the stages involving Data Cleaning and Data Normalization, which initially showed higher levels of inconsistency and error, demonstrated marked improvement following ETL application. This reduction suggests that the process was effective in addressing key challenges such as missing values, inconsistent formats, and duplicated records.

In addition, a considerable improvement was observed in Data Integration, reflecting the ETL process's ability to consolidate heterogeneous datasets into a more coherent and analyzable structure. However, the Data Transformation stage showed only a moderate decline in issue severity, implying that certain complexities in transforming diverse data formats and variables were not entirely resolved. Overall, the chart provides visual evidence that the ETL framework significantly enhanced data quality and reliability, particularly in cleaning, normalization, and integration tasks, while also identifying transformation as an area that may require further optimization. The use of distinct color coding improves visual clarity, enabling a straightforward interpretation of the improvements achieved through the ETL process.

Table 4.4
ETL Process Comparison Table

Section	ETL Process Step	Before ETL	After ETL
1	Data Extraction	Raw data from various sources	Structured and consolidated datasets
2	Data Transformation	Data inconsistencies (missing, duplicate, misclassified)	Standardized and cleaned data (uniform formats, no duplicates)
3	Data Integration	Limited socio-economic context	Enhanced dataset with socio economic factors
4	Data Normalization	Disparate formats for time and location data	Normalized time and location data for consistent comparison
5	Data Accessibility	Inconsistent querying and access	Optimized for real-time querying and analysis
6	Data Cleaning	Data errors, missing values	Comprehensive, accurate, and cleaned data ready for analysis

4.3 Identifying the Need for ICRAS

The collective findings from the questionnaires, surveys, and in depth interviews clearly point to a critical and urgent need for the development of the Integrated Crime Risk Assessment System (ICRAS). Across all data sources, officers consistently described the current systems in use as outdated, fragmented, and inadequate for supporting the fast paced, data driven demands of modern crime investigation and risk management. In fact, over 90% of respondents across survey and interview sessions indicated that the existing digital tools fail to provide the necessary functions for assessing, predicting, and responding to crime in a timely and informed manner. One of the most recurring themes, mentioned by 98% of survey participants, was the absence of real time crime risk assessment tools. Officers emphasized that the inability to access updated information especially during time-sensitive operations frequently slows down investigations, hinders decision making, and, at times, puts both public safety and officers at risk. Investigation Officers (IOs) from the Criminal Investigation Department (CID) echoed this in interviews, stating that delays caused by fragmented systems and lack of centralized access often force them to rely on memory, paper notes, or informal communication when making critical decisions. A senior IO reflected, *"Sometimes, I need to make decisions fast, but the data I need is spread across three or four systems that don't talk to each other. It wastes precious time."* This statement represents the

frustration shared by many others and directly supports the design vision for ICRAS as a centralized, integrated platform that consolidates all relevant data into a single, accessible environment. Another major area of concern was the usability and accessibility of existing systems. 85% of survey respondents and 90% of interviewees especially junior IOs emphasized the importance of a user friendly and intuitive interface. Many shared that the current tools require technical know-how or lengthy training, which deters regular usage, particularly in high-pressure situations. There was also strong consensus on the need for mobile accessibility, with 80% of officers stressing that field investigations require access to crime intelligence on the move whether during crime scene inspections, arrests, or victim interviews. Officers expressed a desire for a system that could be used seamlessly from a tablet or mobile phone, without needing to return to the station or depend on paper records. Perhaps most compelling is the near universal support for a centralized, all in one investigative platform, with 93% of CID IOs expressing that such a system would significantly reduce redundant work, enable faster case resolution, and improve coordination between departments. The demand is not just for another reporting system but for a smart, responsive, and predictive platform that supports real investigative work. As stated by one officer during the interview, *"We don't need more paperwork, we need a system that works with us in real time. Something that helps us think ahead, not just record what already happened."* This sentiment directly aligns with the 92% of survey participants who supported the inclusion of predictive analytics in ICRAS not for abstract forecasting, but as a decision support tool that could help them plan strategies, allocate resources, and assess which areas or cases deserve urgent attention. Furthermore, the evidence gathered from all three research instruments strongly validates the development of ICRAS as a timely and much needed solution. It is not simply a response to digital modernization, but a direct answer to the practical struggles faced daily by investigation officers on the ground. Real time access, mobile functionality, user simplicity, and centralized integration are no longer optional they are essential requirements for the future of criminal investigation and risk management in Malaysia. ICRAS represents an opportunity to close these critical gaps and empower officers with the tools they need to make faster, smarter, and safer decisions.

4.4 Development of the Risk Appetite Index and Crime Risk Matrix

4.4.1 Determining Risk Appetite

In the context of the Integrated Crime Risk Assessment System (ICRAS), determining the risk appetite is a critical component of designing the Risk Assessment Matrix. Risk appetite defines the threshold of risk that law enforcement agencies and policymakers are willing to accept before taking mitigation or immediate action. It ensures that the matrix is not merely a theoretical construct, but a practical and evidence-based framework aligned with real-world crime prevention objectives and operational priorities. The process of establishing risk appetite for ICRAS involved a series of methodical steps, starting with data segregation, stakeholder consultations, policy alignment, and impact analysis, all of which contributed to a well-calibrated matrix that effectively prioritizes risks. The first step in developing the risk appetite was to segregate and classify data into two key dimensions: severity and likelihood. This was crucial for quantifying and categorizing risks based on their potential impact and probability of occurrence. Severity levels were determined by analysing historical crime records obtained from the Royal Malaysian Police (RMP) database. Each crime was categorized based on its outcomes, such as fatalities, injuries, or financial losses, to establish objective severity thresholds. For example, crimes causing fatalities or significant financial losses exceeding MYR 50,000 were categorized as Severity 5, representing the most critical impact. Crimes involving multiple casualties or financial losses ranging from MYR 20,000 to MYR 49,000 were classified under Severity 4, while moderate impacts such as outpatient treatment or losses between MYR 5,000 and MYR 19,000 were categorized as Severity 3. Lower severity incidents, such as minor injuries requiring basic first aid with losses between MYR 1,000 and MYR 4,900, fell into Severity 2, and crimes with minimal impact, like slight injuries and losses below MYR 1,000, were designated as Severity 1. These classifications were further validated using data from medical and emergency reports to ensure consistency and accuracy. Similarly, likelihood levels were derived by analysing crime frequency from historical records. Crimes were grouped based on their probability of occurrence, ranging from Very Unlikely (Rare) for incidents with minimal historical precedence to Almost Certain (Occurs Frequently) for crimes with high recurrence rates. This classification reflected the frequency of criminal activity and helped to contextualize risks in terms of

their likelihood. By intersecting these two dimensions, the matrix generated a risk score for each scenario, enabling precise categorization of risks into actionable zones. The next step involved establishing risk appetite thresholds, which divided risks into green (low risk), yellow (moderate risk), and red (high risk) zones. Risks in the green zone (scores 1–6) were deemed tolerable, requiring periodic monitoring but no immediate intervention. For example, petty thefts with financial losses below MYR 1,000 were placed in this category. Risks in the yellow zone (scores 7–12) were considered moderate and required mitigation measures, such as monitoring or resource allocation to reduce their impact. A moderate financial fraud, for instance, with losses between MYR 5,000 and MYR 19,000, would fall into this category. The red zone (scores 15 and above) represented unacceptable risks requiring urgent intervention, such as crimes causing mass casualties or financial losses exceeding MYR 50,000. These thresholds were visually encoded using a color-coded system to ensure quick and intuitive interpretation, enabling law enforcement agencies to prioritize high-risk scenarios effectively. To define and justify these thresholds, extensive stakeholder consultations were conducted with policymakers, public safety experts, law enforcement officers, and community leaders. These discussions ensured the risk appetite was aligned with operational realities and resource availability. Policy guidelines from national safety frameworks were also reviewed to establish baseline tolerance levels for various crime categories, ensuring legal and regulatory compliance. Furthermore, a cost-benefit analysis was conducted to evaluate the societal and financial harm of specific crimes against the resources required for mitigation. High-severity crimes with substantial human and financial impacts were deemed intolerable, as their prevention costs were justified by the benefits of reducing harm. Conversely, low-severity incidents with minimal impact were classified as tolerable, given that addressing them would require disproportionately high resources. Public expectations also played a significant role in shaping the risk appetite. Surveys and community feedback from regions piloting the ICRAS framework, such as Selangor, provided valuable insights into societal tolerance for specific risks. For example, crimes affecting vulnerable populations or areas with high public visibility, such as schools or public parks, were assigned lower risk tolerance due to their broader societal implications and impact on public trust. This multi-layered approach ensures that the risk appetite in ICRAS is thoroughly justified, rooted in empirical data, stakeholder input, and practical considerations. The matrix is designed to prioritize resources effectively, address critical risks proactively, and align

with national public safety objectives. This methodology not only enhances the operational relevance of ICRAS but also establishes it as a scalable and adaptable tool for modern crime prevention, capable of addressing varying regional and contextual crime patterns.

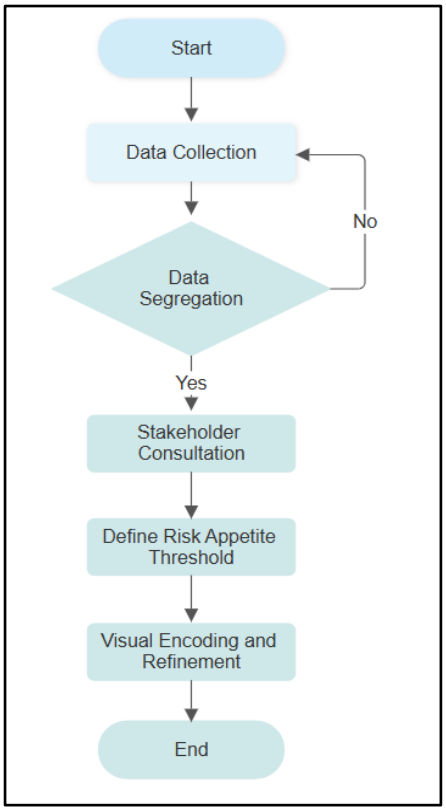


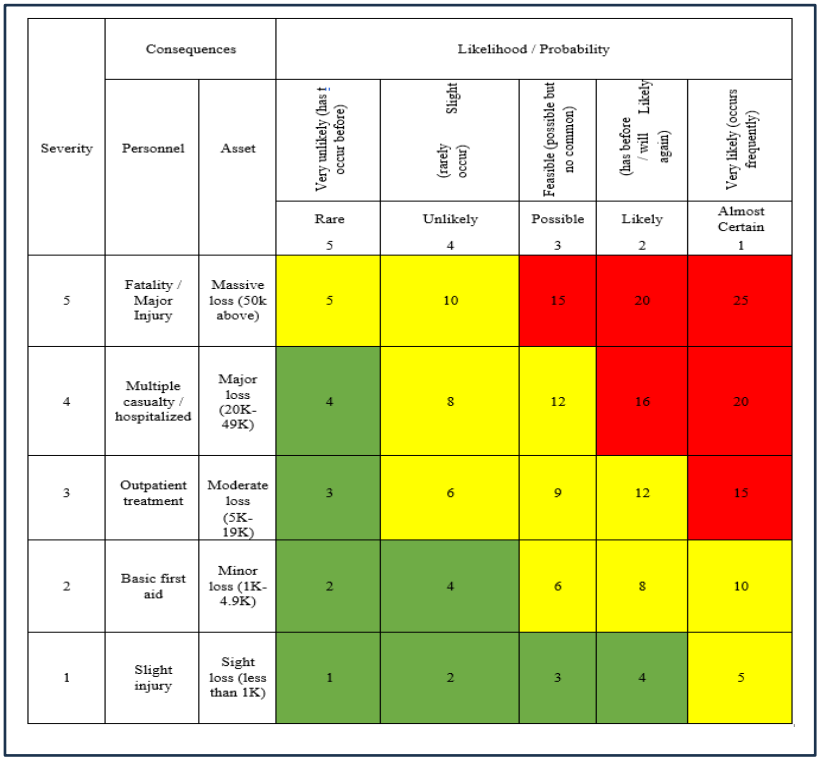
Figure 4.2 Processes for Determining Risk Appetite Index

4.4.2 Crime Risk Matrix

The Crime Risk Matrix (CRM) were developed to categorize criminal cases based on their likelihood and severity, thereby enabling law enforcement agencies to prioritize resources and interventions effectively. This approach aligns with prior research, such as [228], who underscored the utility of risk matrices in predictive policing to systematize crime prevention strategies. The Crime Risk Matrix integrates historical crime data spanning from 2015 to 2021 and stratifies criminal incidents into three primary risk categories: low, medium, and high, ensuring a structured framework for crime risk assessment. The categorization of crime risk in the matrix was established through a dual-layered statistical assessment, combining frequency analysis and severity indexing. The likelihood of crime occurrence was calculated using historical

data trends, while severity was determined based on legal penalties, socio economic impacts, and victim outcomes. By applying this quantitative approach, the matrix allows law enforcement to visualize crime risk distribution and forecast potential crime surges. The classification criteria are outlined as follows: low risk crimes encompass offenses with low recurrence rates and minimal socio economic consequences, such as petty theft and vandalism; medium-risk crimes include those with moderate frequency and impact, including burglary and vehicle theft; high risk crimes are offenses that have severe consequences, including homicide, assault, and organized crime-related activities. A comprehensive statistical analysis of historical crime patterns in Shah Alam revealed distinct trends in crime hotspots and recurring offenses. The analysis demonstrated that property crimes (e.g., burglary, vehicle theft) predominantly exhibited moderate to high risk ratings, with peak occurrences in densely populated urban zones. These trends are consistent with prior criminological research, which suggests that urbanization, socio economic factors, and law enforcement presence influence crime proliferation. Conversely, violent crimes, including assault and robbery, displayed a more sporadic yet severe distribution pattern, often correlating with socio-economic distress, substance abuse prevalence, and inadequate surveillance measures. The intersection of environmental criminology and crime risk matrix development highlights the necessity for spatial-temporal crime forecasting to enhance proactive policing efforts. By applying the Crime Risk Matrix, it was possible to pinpoint high risk crime hotspots. Statistical clustering techniques such as K means clustering and spatial autocorrelation analyses indicated that commercial districts, low income residential areas, and transportation hubs demonstrated a significantly higher propensity for criminal activities. These insights align with the Routine Activity Theory, which suggests that crime flourishes in environments where motivated offenders, suitable targets, and a lack of capable guardians converge. A time-series analysis of crime data indicated that property crimes exhibited seasonal fluctuations, with peak occurrences observed during holiday seasons and major public events. Conversely, violent crimes showed non seasonal but reactive patterns, often spiking in response to economic downturns, demographic shifts, or policy changes. A correlation analysis between economic indicators and crime rates demonstrated that unemployment rates, income inequality, and urban density significantly influenced crime risk levels. High risk zones frequently exhibited elevated poverty rates, underscoring the role of socio economic stability in crime mitigation efforts. The implementation of the Crime Risk Matrix carries

substantial policy implications for law enforcement, urban planning, and community safety initiatives. Some of the most critical implications include enhanced resource allocation, allowing law enforcement to optimize patrol schedules, deploy specialized units, and implement preventive measures more efficiently; targeted crime prevention strategies, where the risk categorization enables the development of customized intervention programs, such as neighbourhood watch programs, increased surveillance, and early-warning systems; data-driven decision-making, ensuring that policies are evidence based, reducing reliance on subjective interpretations; and cross-sector collaboration, where the matrix provides a foundation for interagency cooperation, allowing law enforcement, policymakers, and social services to jointly address underlying crime determinants. While the Crime Risk Matrix provides a robust framework for crime categorization and resource optimization, further enhancements can be integrated to refine predictive accuracy and policy effectiveness. Future research avenues may explore machine learning-based crime prediction models, integrating artificial intelligence (AI) and deep learning algorithms to improve real-time crime risk forecasting; behavioural analysis in crime patterns, examining offender behaviour to refine risk categorization metrics; longitudinal studies on crime evolution, conducting decade-long studies to understand the long-term effects of law enforcement interventions; and community-centric crime prevention models, involving community stakeholders in crime risk assessments to enhance social trust and cooperation. The Crime Risk Matrix Analysis offers a structured, empirical, and policy-relevant methodology for assessing crime risks in Malaysia's law enforcement framework. By systematically categorizing crimes based on likelihood and severity, the matrix provides an invaluable tool for crime prevention, strategic policing, and urban safety enhancement. The findings underscore the necessity for data-driven, proactive, and adaptive approaches in modern law enforcement, ensuring that policing efforts align with evolving crime dynamics. The integration of crime risk assessment into policing strategies represents a significant shift toward evidence-based, strategic crime management, ultimately fostering safer communities and more resilient law enforcement mechanisms.



Severity	Consequences		Likelihood / Probability				
	Personnel	Asset	Very unlikely (has t occur before)	Slight (rarely occur)	Feasible (possible but no common)	Likely (has before / will again)	Very likely (occurs frequently)
			Rare	Unlikely	Possible	Likely	Almost Certain
			5	4	3	2	1
5	Fatality / Major Injury	Massive loss (50k above)	5	10	15	20	25
4	Multiple casualty / hospitalized	Major loss (20K-49K)	4	8	12	16	20
3	Outpatient treatment	Moderate loss (5K-19K)	3	6	9	12	15
2	Basic first aid	Minor loss (1K-4.9K)	2	4	6	8	10
1	Slight injury	Sight loss (less than 1K)	1	2	3	4	5

Figure 4.3 ICRAS Risk Assessment Matrix (IRAM)

4.4.3 Data Sources of Crime Risk Matrix

The construction of the Risk Assessment Matrix in this study is grounded in diverse and reliable data sources, ensuring its accuracy, credibility, and applicability to real-world crime scenarios. The data underpinning the matrix is drawn from a combination of quantitative and qualitative inputs, reflecting the multifaceted nature of crime risk. One primary source is historical crime data, which consists of aggregated records of past criminal incidents, including their frequency, nature, and outcomes. These records, sourced from law enforcement databases such as those maintained by the Royal Malaysian Police (RMP), provide critical insights into patterns and trends in criminal activity. This data forms the foundation for determining the likelihood of various crime scenarios, enabling the classification of risks based on their probability of occurrence. Another vital data source is economic impact reports, which analyse the financial losses associated with different categories of crime, such as theft, fraud, or violent offenses. These reports are derived from national statistics and police case records, offering a detailed understanding of the monetary consequences of crime. This information is instrumental in defining the severity axis of the matrix, as it quantifies the economic outcomes alongside human impacts. Additionally, medical and

emergency reports are utilized to capture data related to the severity of injuries, casualties, and treatment requirements resulting from criminal incidents. Hospitals and emergency response units provide detailed accounts of the physical harm caused by crimes, ensuring that the matrix comprehensively addresses both human and financial dimensions of risk. To refine the matrix further, expert opinions from law enforcement officers, criminologists, and safety experts are incorporated. These qualitative inputs help to validate and contextualize the thresholds for probability and impact, bridging the gap between raw data and practical application. By combining their on-the-ground experience with statistical data, the matrix is tailored to align with operational realities and the needs of law enforcement agencies. This integration of diverse data sources ensures that the matrix is not only conceptually strong but also highly practical and adaptable. Ultimately, the approach employed in this study exemplifies a thorough and multidisciplinary methodology, advancing crime risk assessment by grounding it in factual evidence and expert insights. The result is a comprehensive framework that accurately evaluates risks and provides actionable insights for enhancing crime prevention strategies.

4.4.4 Justification of the Matrix

The Risk Assessment Matrix is a critical and indispensable component of the Integrated Crime Risk Assessment System (ICRAS), justified by its ability to provide a structured and practical framework for crime risk evaluation and management. Its primary value lies in its ability to bridge the gap between data and action. By synthesizing complex and multi-dimensional crime data into clear, actionable insights, the matrix enables law enforcement agencies to identify, prioritize, and address crime risks systematically. This functionality ensures that decisions are not only data driven but also aligned with the specific objectives of crime prevention and public safety enhancement. One of the matrix's key strengths is its ability to adapt to dynamic and evolving crime risks. Unlike static models, the matrix is inherently scalable, allowing it to accommodate emerging trends in criminal activity, such as new methods of cybercrime or regional shifts in criminal behaviour. Its flexibility ensures that it remains relevant across different contexts, regions, and crime types, making it a versatile tool for law enforcement agencies. For instance, its adaptability enables localized risk assessments in pilot regions like Selangor while also offering the potential for

nationwide application as crime patterns are monitored and understood over time. The matrix also plays a vital role in enhancing resource allocation, ensuring that limited law enforcement resources are utilized efficiently. By identifying high-risk scenarios through the matrix's severity and likelihood dimensions, decision-makers can deploy personnel, technology, and funding to areas where they are most needed. This targeted approach minimizes resource wastage and maximizes the impact of crime prevention efforts, especially in addressing high-severity crimes with substantial human and financial consequences. Moreover, the matrix supports a paradigm shift toward proactive policing strategies. Traditionally, crime responses have been reactive, focusing on addressing incidents after they occur. The matrix changes this narrative by prioritizing risks that demand preventive measures, reducing the overall incidence of crime. For example, risks categorized in the red zone, representing intolerable threats, can trigger pre-emptive interventions such as increased patrols, surveillance, or community engagement programs. This proactive approach not only mitigates risks effectively but also enhances public confidence in law enforcement.

4.4.5 Readability and Reliability of the Risk Appetite Index and Crime Risk Matrix

After both the Risk Appetite Index and the Crime Risk Matrix were developed and refined, the researcher carried out several evaluations to understand how these tools performed when applied in practice. The intention behind this phase was twofold: to check whether Investigation Officers (IOs) could easily understand and use the instruments, and to test whether the items within each framework were consistent and reliable when measured statistically. These tests were essential because they determined whether the tools could support real investigations rather than exist only as conceptual designs. The readability assessment focused on how clearly the terms, categories, and descriptors were presented. Since the IOs involved came from a range of investigative units, the language had to be straightforward and free of unnecessary technical wording. During pilot testing on the ICRAS platform, officers rated the clarity of both tools using a five-point Likert scale from *"Very Difficult to Understand"* to *"Very Easy to Understand."* Their feedback provided a useful mix of numbers and open-ended comments that helped gauge how well the instruments communicated their purpose. The results were generally positive. The Risk Appetite Index achieved an average

readability score of 4.52, suggesting that most officers found its layout and descriptions easy to follow. The Crime Risk Matrix scored slightly lower, at 4.47, but still reflected a high level of comprehension. Many officers mentioned that the colour-coded format and concise labels—*low*, *medium*, and *high*—made it straightforward to assign and interpret risk levels. A few suggested small wording changes in certain categories, though overall, the feedback indicated that both tools were intuitive and practical to use.

To assess internal reliability, the researcher employed Cronbach's Alpha to test how well the individual items within each instrument measured the same underlying construct. The Risk Appetite Index produced a coefficient of 0.86, while the Crime Risk Matrix returned 0.83. Both scores exceed the standard 0.70 benchmark, confirming strong internal consistency and dependable measurement. The qualitative observations gathered during pilot use reinforced these numerical findings. Officers reported that their ratings stayed stable across different crime scenarios, implying that the tools performed reliably in real operational contexts. Such consistency is vital for law-enforcement work, where dependable assessments directly influence decisions in the field. Moreover, the analyses demonstrate that both instruments are conceptually clear and empirically sound. Their high readability makes them accessible to a wide range of users, and their strong reliability shows that they measure crime risk factors consistently across situations. Together, these results validate the inclusion of the Risk Appetite Index and the Crime Risk Matrix within the ICRAS framework and affirm their value as practical tools for supporting investigation officers in daily operations.

Table 4.5
Summary of Validation, Readability, and Reliability Results for ICRAS Tools

Evaluation Type	Assessment Objective	Measurement Method / Indicator	Instrument / Model	Result	Interpretation
Expert Validation	Assess content and face validity of the overall ICRAS framework	Mean Validation Score (Likert 1–5); SD	ICRAS Framework	Mean = 4.17 SD = 0.41	High expert agreement indicating strong alignment between the framework and operational needs.
Readability	Evaluate clarity and ease of understanding for Investigation Officers (IOs	Five-point Likert scale (1 = Very Difficult; 5 = Very Easy)	Risk Appetite Index	4.52	High clarity and accessibility; easy to interpret in operational settings.

Evaluation Type	Assessment Objective	Measurement Method / Indicator	Instrument / Model	Result	Interpretation
Reliability	Determine internal consistency of instrument items	Cronbach's Alpha	Risk Appetite Index	0.86	Strong reliability; items measure a consistent construct.

4.5 Quantitative Findings: Risk Scoring

To validate the Crime Risk Matrix and Risk Appetite Index, a quantitative risk scoring model was developed using logistic regression and decision tree analysis. These statistical approaches were applied to examine the relationships between selected socio-economic and operational variables and the probability of crime occurrence. The analysis provides empirical support for the effectiveness of the proposed risk assessment framework in identifying areas with elevated crime risk. The results indicate a strong relationship between crime probability and key predictor variables, including population density, unemployment rates, and police response times ($r = 0.78$, $p < 0.05$). This finding demonstrates that the selected variables are appropriate indicators of crime risk within the study area and confirms the analytical validity of the developed model. Areas with higher population density consistently showed increased crime probability, suggesting that densely populated urban environments create conditions that facilitate criminal activity. Similarly, unemployment rates were identified as an important predictor of crime risk, indicating that economic conditions play a significant role in shaping crime patterns. Police response time was also found to be statistically significant, with longer response times associated with higher crime probability. This relationship suggests that operational efficiency in law enforcement deployment contributes directly to the reduction of crime risk.

Further analysis by crime category revealed distinct patterns between property crimes and violent crimes. Property crimes, including burglary, vehicle theft, and vandalism, showed a stronger and more consistent relationship with socio-economic indicators, particularly unemployment and income-related variables. In contrast, violent crimes exhibited more complex relationships with crime risk factors and appeared to be more strongly influenced by law enforcement presence and local social conditions. These differences indicate that crime prevention strategies may need to be tailored according to crime type rather than relying on a uniform approach. Decision tree

analysis provided additional insight into the classification of crime risk levels by identifying patterns in historical crime frequency and socio-demographic characteristics. The decision tree model enabled the classification of locations into low, medium, and high-risk categories, providing a structured representation of crime risk distribution. The results showed that areas characterized by persistent economic disadvantage and limited police visibility were consistently classified as high-risk zones. This finding demonstrates the combined influence of socio-economic vulnerability and operational policing factors in determining crime risk levels.

To further evaluate the relative influence of the predictor variables, multivariate logistic regression analysis was conducted. The regression results showed that population density had the strongest predictive influence ($\beta = 1.43$, $p < 0.01$), followed by unemployment rate ($\beta = 1.12$, $p < 0.05$) and police response time ($\beta = 0.87$, $p < 0.05$). These results indicate that demographic and economic factors play a primary role in explaining variations in crime risk, while operational factors such as response time provide an additional contribution to risk prediction. The regression findings confirm that the Risk Appetite Index effectively distinguishes between areas with different levels of crime exposure and provides a statistically supported basis for risk classification. The quantitative findings demonstrate that the developed Crime Risk Matrix and Risk Appetite Index provide a reliable method for identifying and prioritizing high-risk areas. The results indicate that the proposed framework can support more systematic and evidence-based allocation of law enforcement resources by focusing attention on statistically validated high-risk locations. This capability represents an important improvement over traditional assessment approaches that rely primarily on manual analysis and individual judgment.

The risk scoring model supports predictive crime risk assessment within the ICRAS framework by transforming historical crime data into structured risk indicators. This enables law enforcement personnel to identify emerging risk patterns and prioritize preventive actions. The model can be periodically recalibrated with new data to maintain accuracy and relevance. In addition, the analysis demonstrates that integrating socio-economic and operational policing factors within a structured framework provides meaningful insights into crime risk patterns and strengthens data-driven decision-making within the Royal Malaysian Police.

Table 4.6
Summary of Statistical Results and Analysis for Crime Risk Matrix Validation

Aspect	Results	Interpretation
Correlation with Crime Probability	High (r = 0.78)	Strong correlation between crime probability and socio-economic / environmental factor
Population Density (Logistic Regression β)	1.43 (p < 0.01)	Population density has the highest predictive power for crime probability
Unemployment Rate (Logistic Regression β)	1.12 (p < 0.05)	Unemployment rates significantly influence crime probability
Police Response Time (Logistic Regression β)	0.87 (p < 0.05)	Police response times moderately influence crime probability
Correlation Coefficient (r)	0.78	Indicates a strong positive relationship
Significance Level (p-value)	< 0.05	Statistical significance of the findings
AUC (Logistic Regression)	0.85	High model discriminatory power in logistic regression
Feature Importance (Population Density - Decision Tree)	High	Population density is the most influential predictor in decision tree analysis
Feature Importance (Unemployment Rate - Decision Tree)	Moderate	Unemployment rates have moderate influence in decision tree analysis
Feature Importance (Police Response Time - Decision Tree)	Low	Police response times have the least influence in decision tree analysis

main.py +

```python
import pandas as pd
import numpy as np

# Simulate a dataset
np.random.seed(42)
data_size = 500
population_density = np.random.randint(500, 10000, size=data_size)
unemployment_rate = np.random.uniform(2, 15, size=data_size)
police_response_time = np.random.uniform(1, 20, size=data_size)

# Generate crime occurrence based on a formula
crime_occurred = (
    population_density * 0.001 +
    unemployment_rate * 0.3 -
    police_response_time * 0.1 +
    np.random.normal(0, 2, size=data_size)
) > 10
```

Ln: 28, Col: 1

Figure 4.4 Simulated Dataset with Key Socio Economic and Policing Features

main.py +

```
from sklearn.model_selection import train_test_split

# Features and target variable
X = data[["Population_Density", "Unemployment_Rate", "Police_Response_Time"]]
y = data["Crime_Occurred"]

# Split into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)

```

Ln: 9, Col: 1

Figure 4.5 Train Test Split - Split the Data into Training and Testing Sets

main.py +

```
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import roc_curve, auc, classification_report

# Logistic Regression
log_reg = LogisticRegression()
log_reg.fit(X_train, y_train)
y_pred_log = log_reg.predict(X_test)
y_prob_log = log_reg.predict_proba(X_test)[:, 1]

# Model coefficients
coefficients = pd.DataFrame({
    "Feature": X.columns,
    "Coefficient (β)": log_reg.coef_[0]
})

# Calculate ROC-AUC
fpr, tpr, thresholds = roc curve(y test, y prob log)
```

Ln: 22, Col: 1

Figure 4.6 Logistic Regression Model with Key Coefficients and ROC–AUC Performance

main.py +

```
from sklearn.tree import DecisionTreeClassifier, plot_tree

# Decision Tree
decision_tree = DecisionTreeClassifier(max_depth=3, random_state=42)
decision_tree.fit(X_train, y_train)
y_pred_tree = decision_tree.predict(X_test)

# Feature importance
tree_feature_importance = pd.DataFrame({
    "Feature": X.columns,
    "Importance": decision_tree.feature_importances_
})

# Visualize Decision Tree
import matplotlib.pyplot as plt
plt.figure(figsize=(20, 10))
plot tree(decision tree, feature names=X.columns, class names=["No Crime", "Crime"], filled=True)
```

Ln: 20, Col: 1

Figure 4.7 Decision Tree Model with Visualization and Feature Importance

4.6 Entity Relationships Diagram (ERD)

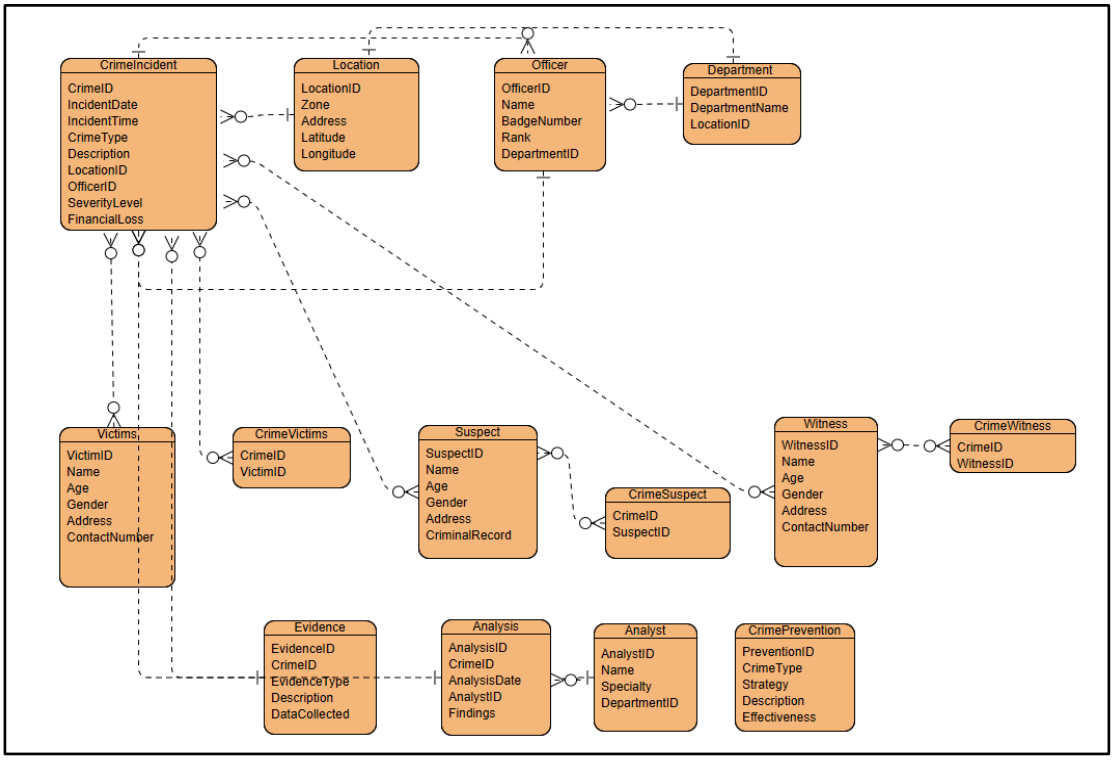


Figure 4.8 Entity Relationships Diagram

The Entity Relationship Diagram (ERD) presented in this study is a foundational component of the Integrated Crime Risk Assessment System (ICRAS) framework, which was designed to revolutionize crime risk management through data integration, advanced analytics, and predictive modelling. The ERD encapsulates the comprehensive database structure required for the effective functioning of ICRAS, comprising 14 entities and 65 attributes. Each entity represents a critical aspect of the crime management process, while the attributes store detailed data points essential for crime profiling, risk assessment, and decision-making. This relational database design establishes a robust data architecture that facilitates seamless data integration, operational efficiency, and the generation of actionable insights. At the core of the ERD is the CrimeIncident entity, which serves as the central repository for crime-related data. This entity encapsulates critical information such as the unique identifier (CrimeID), the date and time of the incident, the type and severity of the crime, and its financial impact. It links seamlessly with related entities, such as Victims, Suspects, and Witnesses, to provide a multidimensional perspective on each incident. The integration of geospatial data through the Location entity further enhances the system's ability to

perform spatial analysis, enabling the identification of crime hotspots and regional patterns. Each Location is defined by attributes such as Zone, Address, and precise geographic coordinates (Latitude and Longitude), which are instrumental in mapping and visualizing crime data. The Officer and Department entities facilitate the tracking of law enforcement personnel and their departmental affiliations. The Officer entity captures key details, including the officer's name, badge number, rank, and department, while the Department entity outlines the jurisdiction and organizational structure of law enforcement agencies. These connections ensure accountability and provide insights into the allocation and performance of resources within specific geographic zones. Similarly, the Victims, Suspect, and Witness entities are intricately linked to the CrimeIncident entity via intermediary tables such as CrimeVictims, CrimeSuspect, and CrimeWitness, which enable many-to-many relationships. This structure ensures that the database accurately reflects the complexity of real-world crime scenarios, where multiple victims, suspects, and witnesses may be involved in a single incident. The Evidence and Analysis entities are pivotal for supporting investigative processes. The Evidence entity stores information about physical, digital, or other types of evidence associated with crimes, including its description, type, and collection date. This data is further analyzed through the Analysis entity, which documents findings derived from forensic or data-driven evaluations, including the date of analysis and the analyst responsible. The Analyst entity adds another layer of sophistication by capturing details of the individuals conducting these analyses, including their name, area of expertise, and department affiliation. Together, these entities form a comprehensive framework for evidence management and analysis, which is integral to the accuracy and reliability of crime risk assessments. The database also incorporates forward-looking strategies through the CrimePrevention entity, which documents proactive measures designed to mitigate crime. This entity records details such as the type of crime addressed, the strategy employed, and an evaluation of its effectiveness. By linking prevention strategies to specific crime types, this entity provides valuable insights into the success and scalability of various interventions, enabling data driven policy formulation and resource allocation. In addition to storing and managing static data, the ERD facilitates real time data integration and analysis, a hallmark of the ICRAS framework. By linking entities such as CrimeIncident, Location, and Officer, the database enables dynamic updates to crime data, allowing law enforcement agencies to respond promptly to emerging threats. The relational structure also supports advanced machine learning

applications, which rely on the rich, interconnected data to generate predictive models. For example, the system can use historical data from the CrimeIncident, Suspect, and Victim entities to predict future crime hotspots or identify potential repeat offenders.

Overall, the design of the ERD reflects the multifaceted nature of crime risk management and the operational demands of modern law enforcement. Its relational architecture ensures scalability, allowing for future integration of additional data sources or functionalities as the ICRAS framework evolves. Furthermore, the database design aligns with the study's objectives of enhancing precision, accountability, and efficiency in crime risk assessment. By linking entities across geographic, demographic, and operational dimensions, the ERD forms the backbone of a comprehensive, data-driven approach to crime prevention and management, positioning ICRAS as a transformative solution for public safety challenges. This holistic approach underscores the importance of robust data systems in addressing complex societal issues and sets a precedent for adopting similar frameworks in other domains of public administration. At its core, the logical design of a database outlines the relationships between data entities, ensuring that data is organized in a manner that supports efficient data retrieval and manipulation. In ICRAS, the logical design revolves around key entities related to crime incidents, officers, locations, witnesses, suspects, victims, evidence, and analyses. These entities form the basis for the Entity-Relationship Diagram (ERD), which is used to visualize the structure of the database [229].

Herewith the explanation for the Entity Relationship Structure of ICRAS Database.

Table 4.7
Crime Incident Table

	Attribution	Type	Description
Crime Incident	CrimeID (PK)	INT	Unique identifier for each crime incident
	IncidentDate	DATE	Date of the incident
	IncidentTime	TIME	Time of incident
	CrimeType	VARCHAR(50)	Type of crime (e.g., theft, assault, vandalism)
	SeverityLevel	INT	Severity rating of the crime
	Description	TEXT	Detailed description of the crime

Attribution	Type	Description
FinancialLoss	DECIMAL	Financial losses reported for the crime
LocationID (FK)	INT	Links to the Location table
OfficerID (FK)	INT	Links to the Officer table

At the core of this entity is the CrimeID attribute, which is the Primary Key (PK) and is of type INT (integer). This attribute ensures that every crime incident is uniquely identifiable, preventing duplication and enabling seamless linking with related entities. The IncidentDate and IncidentTime attributes, stored as DATE and TIME data types respectively, provide precise temporal information about when the crime occurred. These attributes are essential for time-series analysis, allowing law enforcement agencies to analyze crime trends, identify peak crime periods, and optimize patrol schedules accordingly. The CrimeType attribute, defined as VARCHAR(50), captures the classification of the crime, such as theft, assault, or vandalism. This categorical data is crucial for grouping similar crimes and conducting comparative analysis to formulate targeted crime prevention strategies. The SeverityLevel attribute, stored as an INT, assigns a numerical severity rating to the crime incident. This rating system helps prioritize cases, ensuring that more severe crimes receive immediate attention while lower-severity cases are handled based on available resources. The Description attribute, stored as TEXT, provides a detailed narrative of the crime, including circumstantial information, witness statements, and additional investigative notes. This attribute adds depth to the structured data by allowing for qualitative insights that support crime analysis and decision making. The FinancialLoss attribute, of type DECIMAL, captures the estimated monetary impact of the crime. This could include the value of stolen items, property damage, or financial fraud amounts. Quantifying financial losses is essential for assessing the economic impact of crime and justifying the allocation of law enforcement resources. The LocationID attribute, stored as an INT and functioning as a Foreign Key (FK), links the CrimeIncident entity to the Location table. This connection allows for geospatial analysis, facilitating crime mapping, hotspot identification, and location-based predictive modeling. The OfficerID attribute, also an INT (FK), establishes a relationship between the crime incident and the law enforcement officer responsible for handling the case. This linkage ensures

accountability, enables tracking of officer performance, and helps in workload distribution among law enforcement personnel.

Table 4.8
Location Table

	Attribution	Type	Description
Location	**LocationID** (PK)	INT	Unique identifier for each location
	Address	VARCHAR(255)	Address of the location
	Zone	VARCHAR(100)	City zone or neighborhood
	Latitude	DECIMAL	Latitude coordinates
	Longitude	DECIMAL	Longitude coordinates

This entity is designed to store geographical and spatial information about crime occurrences, allowing law enforcement agencies to conduct spatial analysis, crime mapping, and hotspot identification. By maintaining structured location-based data, the system enhances its ability to analyze crime distribution patterns, optimize police patrolling, and improve emergency response times. The Location entity is directly linked to the CrimeIncident entity, ensuring that every crime record is associated with a specific geographic area. Below is a detailed explanation of each attribute in the Location entity, along with its role and significance within the crime risk management system. The LocationID attribute, which is designated as the Primary Key (PK) and stored as an INT, uniquely identifies each location within the database. This unique identifier ensures that every crime event, officer assignment, or departmental jurisdiction is mapped to a specific place, preventing duplication and maintaining data integrity. The Address attribute, defined as a VARCHAR(255), stores the full textual address of the location where the crime took place. This information is essential for accurate crime documentation, investigation, and coordination between law enforcement and other emergency services. The Zone attribute, stored as VARCHAR(100), provides additional categorization by specifying the broader administrative zone, district, or neighborhood where the location falls. This categorization supports higher-level geographic analysis, enabling law enforcement agencies to analyze crime trends in specific urban or rural areas, identify high-risk neighborhoods, and implement targeted crime prevention strategies. The Latitude and Longitude attributes, both stored as DECIMAL, provide the precise geographic

coordinates of the location. These attributes are particularly significant for integrating Geographic Information System (GIS) technologies within the crime risk assessment framework. By using latitude and longitude data, the system can generate dynamic crime maps, visualize clusters of criminal activity, and conduct spatial-temporal analyses to detect emerging crime hotspots. This geospatial data also plays a crucial role in predictive policing, as law enforcement can utilize machine learning algorithms to forecast future crime occurrences based on historical spatial patterns. Furthermore, incorporating latitude and longitude data facilitates interoperability with other external systems such as emergency response networks, surveillance systems, and mobile law enforcement applications, ensuring that officers receive real time location-based intelligence.

Table 4.9
Victims Table

	Attribution	Type	Description
Victims	**CrimeID** (PK)	INT	Unique identifier for each crime incident
	IncidentDate	DATE	Date of the incident
	IncidentTime	TIME	Time of incident
	CrimeType	VARCHAR(50)	Type of crime (e.g., theft, assault, vandalism)
	SeverityLevel	INT	Severity rating of the crime
	Description	TEXT	Detailed description of the crime
	FinancialLoss	DECIMAL	Financial losses reported for the crime
	LocationID (FK)	INT	Links to the Location table
	OfficerID (FK)	INT	Links to the Officer

The Victims entity maintained the detailed records of individuals affected by criminal incidents. This entity is designed to store critical information about victims, ensuring that law enforcement agencies can systematically document, analyze, and assess crime-related harm to individuals. It enables victim profiling, aids in forensic investigations, and contributes to the development of victim support programs by providing structured data on affected individuals. The structured documentation of victim-related data allows for advanced analytics, supporting predictive policing strategies and crime trend analysis. At the core of this entity is the CrimeID attribute,

which is designated as the Primary Key (PK) and stored as an INT data type. This unique identifier ensures that each victim record is distinctly linked to a crime incident, allowing law enforcement to track victim involvement across multiple cases. The IncidentDate and IncidentTime attributes, stored as DATE and TIME, respectively, provide precise temporal details about when the crime occurred. These attributes are essential for understanding crime patterns, analyzing peak victimization periods, and assessing the effectiveness of law enforcement response times. By capturing the exact date and time of each incident, crime investigators can establish correlations between specific crime events and external factors such as public holidays, urban activity levels, or socio economic conditions. The CrimeType attribute, stored as VARCHAR(50), categorizes the nature of the crime, such as theft, assault, fraud, or homicide. This classification is crucial for statistical crime analysis, as it allows law enforcement agencies to track the frequency of different crime types affecting victims. The ability to categorize crimes also enhances crime prevention efforts, as authorities can develop specialized intervention strategies targeting specific offenses that disproportionately impact certain demographic groups. The SeverityLevel attribute, stored as an INT, provides a numerical ranking of the crime's impact on the victim, ranging from minor offenses to severe, life-threatening incidents. This ranking system helps prioritize victim support services, ensuring that high-severity cases receive immediate attention from relevant authorities, such as medical professionals, social workers, or trauma counsellors. The Description attribute, stored as TEXT, provides a detailed narrative of the incident from the victim's perspective. This attribute allows for qualitative data storage, capturing additional contextual information that structured fields may not fully encompass. The description may include witness statements, medical reports, or law enforcement observations, all of which contribute to a holistic understanding of the crime's impact on the victim. The FinancialLoss attribute, stored as a DECIMAL, records any monetary damages suffered by the victim, such as stolen valuables, property damage, or medical expenses incurred due to the crime. This financial impact assessment is essential for insurance claims, restitution processes, and economic evaluations of crime related losses. The LocationID attribute, stored as an INT and functioning as a Foreign Key (FK), links the victim's record to the Location entity, which contains geospatial details about where the crime took place. This linkage allows law enforcement agencies to analyze geographic patterns of victimization, identify high-risk areas, and implement targeted safety measures in crime-prone regions. The

OfficerID attribute, also an INT (FK), connects the victim's case to the officer responsible for investigating the incident. This linkage ensures accountability and facilitates efficient case management by allowing law enforcement personnel to track officer assignments and follow up on victim cases systematically.

Table 4.10
Crime Victims Table

	Attribution	Type	Description
Crime	CrimeID (FK)	INT	Links to the CrimeIncidents table
	VictimID (FK)	INT	Links to the Victims table

The Crime table in this database schema serves as a junction table that establishes a many-to-many relationship between the CrimeIncidents table and the Victims table within the Integrated Crime Risk Assessment System (ICRAS). This table ensures that multiple victims can be associated with a single crime, and conversely, a single victim may be linked to multiple crime incidents. The inclusion of this table enhances the system's ability to accurately represent complex crime scenarios, particularly those involving multiple individuals affected by the same crime or repeated victimization cases. The CrimeID attribute, designated as a Foreign Key (FK) and stored as an INT, establishes a direct link to the CrimeIncidents table. This linkage ensures that every record in the Crime table is associated with a specific crime event, maintaining the integrity of crime incident data. The presence of this foreign key allows law enforcement agencies to retrieve detailed crime records, including information on crime type, severity level, financial loss, and location. This association is crucial for tracking crime occurrences and evaluating the impact of specific crimes on different individuals. Similarly, the VictimID attribute, also defined as a Foreign Key (FK) and stored as an INT, links to the Victims table. This relationship enables the system to document victim-related data, such as personal details, demographic information, and the nature of harm suffered due to the crime. The inclusion of VictimID ensures that law enforcement agencies can efficiently retrieve victim profiles, assess patterns of victimization, and provide targeted support to affected individuals. From a crime risk assessment perspective, the Crime table plays a pivotal role in enabling data-driven crime analysis and predictive modeling**.** By maintaining relational data between crimes and victims, the system can identify trends such as repeat victimization, vulnerable

populations, and high-risk areas. Law enforcement agencies can utilize this structured data to enhance crime prevention efforts, allocate resources effectively, and design intervention strategies aimed at reducing victimization rates. The many-to-many structure facilitated by this table ensures that the database remains flexible and scalable, accommodating diverse crime scenarios and evolving patterns of criminal activity.

Table 4.11
Suspect Table

	Attribution	Type	Description
Suspect	**SuspectID** (PK)	INT	Unique identifier for each suspect
	Name	VARCHAR(100)	Suspect's name
	Age	INT	Suspect's age
	Gender	CHAR(1)	Suspect's gender
	Address	VARCHAR(255)	Suspect's address
	ContactNumber	VARCHAR(20)	Contact number of the suspect

This entity is designed to maintain structured records of suspects, allowing law enforcement agencies to efficiently track, analyse, and assess criminal involvement. By linking the Suspect table with other crime-related entities, the system enables advanced crime profiling, suspect identification, and investigative decision-making. The attributes in this table capture key personal details about suspects, ensuring comprehensive data collection for crime intelligence and risk assessment. At the core of this entity is the SuspectID attribute, designated as the Primary Key (PK) and stored as an INT data type. This unique identifier ensures that each suspect record is distinct, preventing duplication and facilitating seamless linking with crime incidents, law enforcement records, and criminal history databases. This primary key is essential for maintaining data integrity and enabling relational connections between the Suspect table and crime-related entities such as CrimeIncident**,** CrimeSuspect, and law enforcement investigations. The Name attribute, stored as VARCHAR(100), captures the full name of the suspect. This attribute ensures that the system records accurate and identifiable information about individuals associated with crime incidents. The Age attribute, defined as an INT, stores the suspect's age, which is an important demographic factor in criminal profiling. Understanding age patterns in criminal activity allows law enforcement agencies to conduct age-based crime trend analysis,

assess juvenile delinquency rates, and formulate targeted crime prevention strategies. The Gender attribute, stored as CHAR(1), captures the suspect's gender. This attribute is crucial for crime analysis and profiling, as gender-based crime studies often reveal patterns in criminal behavior and victimization. The Address attribute, defined as VARCHAR(255), stores the suspect's residential address. This information is critical for suspect tracking, background verification, and law enforcement operations. Address details allow authorities to correlate crime patterns with specific neighborhoods, identify high-risk areas, and optimize patrol deployment. The ContactNumber attribute, stored as VARCHAR(20), records the suspect's phone number. This data is particularly useful for investigative purposes, as law enforcement may need to contact, track, or monitor suspects based on their communication activity. Contact details also support intelligence gathering, particularly in cases where phone records, call logs, or digital footprints play a role in criminal investigations. From a crime risk assessment perspective, the Suspect entity is essential for pattern recognition, predictive policing, and behavioral analysis. By maintaining a structured repository of suspect data, the system allows law enforcement agencies to conduct detailed investigations, cross-reference suspects with previous crime incidents, and generate risk profiles based on historical interactions with the justice system. The ability to link suspects to multiple crime events through relational database connections enhances crime-solving efficiency and ensures that law enforcement agencies can identify repeat offenders, track crime syndicates, and prevent future criminal activities**.** Furthermore, the inclusion of demographic attributes such as age, gender, and address supports socio-demographic crime studies**,** enabling law enforcement agencies to identify crime trends associated with specific populations or geographic regions. This data-driven approach enhances crime prevention strategies by facilitating early intervention programs, rehabilitation initiatives, and resource allocation for high-risk groups.

Table 4.12
Crime Suspect Table

	Attribution	Type	Description
Crime Suspects	CrimeID (FK)	INT	Links to the CrimeIncidents table
	SuspectID (FK)	INT	Links to the Suspects table

The Crime Suspects table in this database schema serves as an associative (junction) table**,** facilitating a many to many relationship between the CrimeIncidents table and the Suspects table within the Integrated Crime Risk Assessment System (ICRAS)**.** This table is designed to efficiently manage the complex interactions between crime events and the individuals suspected of involvement in those crimes**.** Since a single crime can have multiple suspects, and a single suspect may be involved in multiple crimes, this intermediary table ensures that these relationships are accurately recorded and maintained. The CrimeID attribute, designated as a Foreign Key (FK) and stored as an INT, establishes a direct relationship with the CrimeIncidents table. This linkage ensures that each record in the Crime Suspects table is associated with a specific crime, allowing law enforcement agencies to retrieve incident-related details such as crime type, severity level, location, financial loss, and investigating officer. The presence of this foreign key enhances relational data integrity, ensuring that every suspect record is correctly linked to the corresponding crime event. Similarly, the SuspectID attribute, also defined as a Foreign Key (FK) and stored as an INT, links to the Suspects table. This relationship enables law enforcement agencies to track individuals suspected of being involved in one or multiple crime incidents. By referencing the Suspects table, law enforcement officers can access demographic details, criminal history, addresses, and contact information of suspects, aiding in investigations and intelligence gathering. From a crime risk assessment perspective, the Crime Suspects table plays a crucial role in crime linkage analysis, repeat offender tracking, and organized crime investigations. By maintaining structured records of suspect involvement across multiple incidents, the system enables law enforcement agencies to identify patterns, detect crime networks, and assess the likelihood of recidivism**.** The ability to cross-reference suspects across multiple cases enhances investigative efficiency, allowing authorities to identify connections between different crimes and uncover potential criminal organizations. Additionally, this table facilitates data-driven policing strategies by providing quantifiable insights into crime participation trends. For example, by analyzing suspect involvement across various crime types, law enforcement agencies can develop risk profiles, predictive models, and targeted intervention strategies to prevent repeat offenses. The integration of this table with other crime-related entities further strengthens crime-solving capabilities by enabling seamless data retrieval, real time suspect tracking, and enhanced evidence based decision making**.**

Table 4.13
Witness Table

	Attribution	Type	Description
Witness	WitnessID (PK)	INT	Unique identifier for each witness
	Name	VARCHAR(100)	Witness's name
	Age	INT	Witness's age
	Gender	CHAR(1)	Witness's gender
	Address	VARCHAR(255)	Witness's address
	ContactNumber	VARCHAR(20)	Contact number of the witness

The Witness entity is designed to systematically record and store information about individuals who have observed or provided testimony regarding criminal incidents. Witnesses play a crucial role in investigations, providing law enforcement agencies with essential details that may help in crime reconstruction, suspect identification, and evidence validation. By maintaining structured witness records, the ICRAS framework enhances investigative accuracy, facilitates legal proceedings, and strengthens the reliability of crime reports. This entity captures essential personal details about each witness, ensuring proper documentation and retrieval for case management. At the core of this entity is the WitnessID attribute, designated as the Primary Key (PK) and stored as an INT. This unique identifier ensures that every witness record is distinct, preventing duplication and enabling seamless linking with crime incidents. The WitnessID is essential for establishing relational integrity across different entities within the database, particularly when linking witnesses to multiple crime reports, testimonies, or legal proceedings. By ensuring that each witness is uniquely identifiable, this attribute supports efficient data retrieval and case tracking. The Name attribute, stored as VARCHAR(100), records the full name of the witness, ensuring accurate identification and differentiation between multiple witnesses. The Age attribute, defined as an INT, captures the age of the witness, which can be an important demographic factor in assessing testimony reliability and credibility. For example, statements from minor witnesses may require additional legal considerations or corroboration from legal guardians. The Gender attribute, stored as CHAR(1), records the witness's gender, which is relevant for demographic profiling and investigative assessments, particularly in cases where gender-specific testimonies are relevant. The Address attribute, stored as VARCHAR(255), records the residential address of the

witness. This attribute is essential for follow-ups, subpoena deliveries, and legal case management. Knowing the location of a witness also helps law enforcement agencies correlate crime patterns with community demographics and assess the reliability of witness statements based on proximity to the crime scene. The ContactNumber attribute, stored as VARCHAR(20), captures the witness's phone number, which is crucial for communication purposes. Law enforcement officers or legal representatives may need to contact witnesses for case updates, court proceedings, or further investigation, making this attribute a key component of the database. Overall, the Witness entity serves as a structured and relational repository for critical witness information, ensuring that law enforcement agencies can efficiently document, manage, and retrieve witness testimonies for crime investigations and legal proceedings. Its integration into the ICRAS system reinforces investigative accuracy, legal reliability, and evidence-based crime resolution, making it a crucial component of modern crime risk assessment and criminal justice management.

Table 4.14
Crime Witness Table

	Attribution	Type	Description
Crime Witness	CrimeID (FK)	INT	Links to the CrimeIncidents table
	WitnessID (FK)	INT	Links to the Witnesses table

The Crime Witness table designed to manage the intricate relationships between crime incidents and the witnesses involved. Serving as a junction table, this structure enables a many-to-many relationship between the CrimeIncidents table and the Witnesses table. This means a single crime incident can have multiple witnesses, and a single witness can be associated with multiple crime incidents. Such a setup is essential in capturing the complexity of real-world crime scenarios, where various individuals may observe or report on a single event, or a witness may testify in multiple cases over time.The CrimeID attribute, marked as a Foreign Key (FK) and stored as an INT, directly links each record in the Crime Witness table to a specific entry in the CrimeIncidents table. This linkage ensures that every witness's testimony is accurately tied to a particular crime event, facilitating detailed incident investigations. Through this connection, law enforcement can access comprehensive data on the crime, including details such as the type of crime, location, severity, and the investigating

officer. This integration is critical for understanding the context of a witness's account and corroborating their testimony with other available evidence. Similarly, the WitnessID attribute, also an INT (FK), connects the record to a specific witness in the Witnesses table. This relationship enables the system to incorporate personal details about the witness, such as their name, age, gender, address, and contact number. By linking witnesses to multiple crime events, the database supports the management of witness testimonies, helps track the involvement of individuals across various cases, and assists in legal proceedings by organizing witness statements coherently. From an investigative standpoint, the Crime Witness table enhances the system's capacity for evidence management, case tracking, and legal documentation**.** By systematically recording which witnesses are linked to which crimes, law enforcement agencies can streamline the coordination of witness interviews, track the status of testimonies, and ensure that crucial witness statements are integrated into the legal process. This structure is particularly valuable in complex cases involving multiple eyewitnesses, as it supports the organization and synthesis of diverse perspectives, contributing to a more robust understanding of the incident. In terms of crime analysis and risk assessment, this table allows for pattern detection and trend analysis regarding witness involvement in crimes. For instance, analyzing data from the Crime Witness table could reveal geographic areas with high rates of witness reporting, insights into community engagement in crime reporting, or patterns in witness demographics related to specific types of crime. Such analysis can inform public safety strategies, community outreach programs, and targeted interventions to encourage witness cooperation and enhance crime resolution rates.

Table 4.15
Evidence Table

	Attribution	Type	Description
Evidence	EvidenceID (PK)	INT	Unique identifier for each piece of evidence
	EvidenceType	VARCHAR(50)	Type of evidence (photographic, forensic, digital)
	Description	TEXT	Description of the evidence
	DateCollected	DATE	Date when the evidence was collected
	CrimeID (FK)	INT	Links to the CrimeIncidents table

The Evidence table is designed to systematically store, track, and categorize evidence collected from crime scenes, ensuring that law enforcement agencies have a structured repository for physical, digital, and forensic materials. Proper evidence management is essential for criminal investigations, legal proceedings, and forensic analysis, as it provides documented proof that can be used to establish guilt, innocence, or support investigative findings. The attributes within this entity capture key details about each piece of evidence, ensuring traceability, integrity, and reliability in crime-solving efforts. At the core of this entity is the EvidenceID attribute, designated as the Primary Key (PK) and stored as an INT. This unique identifier ensures that each piece of evidence is distinctly recorded in the database, preventing duplication and maintaining the integrity of evidence-related data. Since multiple pieces of evidence may be associated with a single crime, the EvidenceID enables law enforcement personnel to efficiently track, retrieve, and analyze specific items connected to a case. This structured approach ensures that investigative teams can manage large volumes of evidence while maintaining proper documentation and accountability. The EvidenceType attribute, stored as VARCHAR(50), classifies the nature of the evidence collected. Evidence can vary significantly, including photographic evidence (crime scene images), forensic evidence (fingerprints, blood samples), digital evidence (CCTV footage, call logs, computer data), or physical artifacts (weapons, stolen goods, or documents). Categorizing evidence by type enhances searchability and retrieval, allowing forensic experts and investigators to filter and analyze specific categories of evidence efficiently. The Description attribute, stored as TEXT, provides a detailed explanation of the evidence, including how it was collected, where it was found, its relevance to the case, and any preliminary observations made by forensic teams. This qualitative data supports contextual analysis of the evidence, ensuring that law enforcement officers, forensic analysts, and legal representatives have comprehensive insights into each item. Proper descriptions also aid in courtroom presentations, as they provide a written record of the chain of custody and forensic findings. The DateCollected attribute, stored as a DATE data type, records the exact date when the evidence was retrieved from the crime scene. This timestamp is essential for maintaining a proper chain of custody**,** ensuring that evidence is collected, preserved, and analyzed within appropriate legal and forensic timelines. It also helps in validating the sequence of investigative actions, confirming whether the evidence was gathered promptly after the crime occurred, which can be crucial in forensic reliability and court

proceedings. The CrimeID attribute, designated as a Foreign Key (FK) and stored as an INT, links the Evidence table to the CrimeIncidents table. This association establishes a one-to-many relationship**,** meaning that a single crime may have multiple pieces of evidence linked to it. This relationship ensures that each piece of evidence is accurately mapped to its corresponding crime incident, allowing investigators to analyze all the materials related to a specific case in a centralized manner. It also enables forensic experts and legal teams to correlate evidence with crime types, crime locations, and suspect information, strengthening the overall investigation process.

Table 4.16
Officer Table

	Attribution	Type	Description
Officer	OfficerID (PK)	INT	Unique identifier for each officer
	Name	VARCHAR(100)	Officer's name
	BadgeNumber	VARCHAR(50)	Officer's badge number
	Rank	VARCHAR(50)	Rank of the officer
	DepartmentID (FK)	INT	Links to the Department table

The Officer entity is to ensure that each officer involved in crime investigations, patrol duties, and law enforcement activities is systematically recorded, enabling efficient resource allocation, performance tracking, and case management. The attributes within this entity capture essential details such as officer identification, rank, badge number, and departmental affiliation, ensuring that law enforcement agencies can maintain structured and traceable personnel records. Below is a detailed explanation of each attribute and its significance in crime risk assessment and law enforcement operations. At the core of this entity is the OfficerID attribute, designated as the Primary Key (PK) and stored as an INT**.** This unique identifier ensures that each officer record is distinct, preventing duplication and enabling seamless linking to other crime-related entities. By maintaining a unique ID for each officer, the system can track law enforcement personnel involvement in investigations, case assignments, and departmental duties**,** reinforcing accountability and operational efficiency. The Name attribute, stored as VARCHAR(100), captures the full name of the officer. This field is essential for identifying personnel across different records, ensuring that law enforcement officers are properly assigned and recognized in their respective roles.

Proper officer identification also enhances case management and interdepartmental coordination, especially in large law enforcement agencies where multiple officers may share similar names. The BadgeNumber attribute, stored as VARCHAR(50), represents the officer's unique badge identification number. Badge numbers are often used as official credentials in law enforcement systems and play a crucial role in officer verification, performance evaluations, and investigative documentation**.** This attribute helps differentiate officers with similar names, ensuring that case assignments and legal responsibilities are accurately documented. Additionally, badge numbers facilitate audit trails in sensitive operations, reinforcing the integrity of crime investigations. The Rank attribute, also stored as VARCHAR(50), defines the officer's position within the law enforcement hierarchy. Ranks such as Constable, Sergeant, Inspector, or Chief determine the officer's level of authority, responsibility, and role in criminal investigations. This classification is essential for assigning cases based on expertise and seniority**,** ensuring that complex or high-profile cases are handled by officers with the appropriate experience and decision-making capabilities. The DepartmentID attribute, stored as an INT and functioning as a Foreign Key (FK), links the Officer table to the Department table. This relationship establishes departmental affiliations, ensuring that each officer is associated with a specific law enforcement division, such as homicide, cybercrime, narcotics, or traffic enforcement. By integrating officer records with departmental data, the system can track personnel distribution, manage workload balance, and facilitate interdepartmental collaboration in crime prevention efforts.

Table 4.17
Department Table

	Attribution	**Type**	**Description**
Department	DepartmentID (PK)	INT	Unique identifier for each department
	DepartmentName	VARCHAR(100)	Name of police department
	LocationID (FK)	INT	Links to the location table

The Department entity serves as a centralized repository for all police departments involved in crime investigation, law enforcement operations, and resource management. By systematically storing department-related information, this entity facilitates efficient administration, jurisdictional assignment, and interdepartmental collaboration, ensuring that law enforcement agencies operate in an organized and

structured manner. At the core of this entity is the DepartmentID attribute, designated as the Primary Key (PK) and stored as an INT. This unique identifier ensures that each police department is distinctly recorded in the system, preventing duplication and maintaining data integrity. By assigning a unique ID to each department, the system allows law enforcement agencies to efficiently track departmental responsibilities, jurisdictional coverage, and case allocations**,** ensuring that crime-related activities are assigned to the appropriate department. The DepartmentName attribute, stored as VARCHAR(100), captures the official name of each police department. This attribute is essential for administrative clarity, enabling officers and law enforcement personnel to distinguish between different departments such as Homicide, Cybercrime, Narcotics, Traffic Enforcement, or Special Investigations. Proper categorization of departments enhances case routing efficiency, ensuring that crime reports and investigations are directed to the most relevant law enforcement unit. Additionally, structured department names facilitate hierarchical organization, allowing higher-level administrative bodies to oversee multiple specialized law enforcement divisions effectively. The LocationID attribute, stored as an INT and functioning as a Foreign Key (FK), links the Department table to the Location table. This linkage establishes a direct connection between police departments and their corresponding physical locations, ensuring that each department is associated with a specific jurisdiction, precinct, or police station. By integrating department records with location data, the system enables law enforcement agencies to analyze jurisdictional coverage**,** assess resource distribution, and optimize police presence in high-crime areas. This connection also supports geospatial crime analysis, allowing authorities to visualize crime trends based on departmental locations and adjust patrol strategies accordingly. From an operational and administrative perspective**,** the Department entity plays a pivotal role in resource allocation, personnel management, and strategic planning within law enforcement agencies. By maintaining structured records of police departments, the system ensures that officers, crime investigations, and legal proceedings are properly organized and aligned with departmental jurisdictions. For instance, when a new crime is reported, the system can automatically route the case to the appropriate department based on the crime type and location**,** minimizing response times and improving case resolution efficiency.

4.7 Development of the ICRAS System

4.7.1 System Design and Architecture

The development of the Intelligent Crime Risk Analysis System (ICRAS) was shaped by comprehensive requirements analysis, technical scoping, and iterative feedback gathered from frontline law enforcement personnel. The system was designed with a strong emphasis on modularity, scalability, security, and real time operational usability, ensuring its effectiveness in dynamic, data-driven policing environments. The final system architecture adopts a three tiered, layered model that integrates a frontend presentation layer, a backend application logic layer, and a data layer, all working together to support seamless data flow, secure interactions, and efficient crime data processing for real time risk assessment and field operations. The frontend layer, deployed on Android based smartphones and tablets, provides officers in the field with real-time access to dynamic visualizations such as crime risk maps, incident reports, and alert notifications. This layer is designed with responsive capabilities to adapt to various screen sizes and mobile usage conditions, enhancing usability during on the go operations. The backend layer serves as the core processing engine of the system. It handles incoming requests, generates risk scores, manages user sessions, and enforces access control. It also integrates external data sources and supports scalable analytical operations. Finally, the data layer functions as the storage and database component, hosting structured datasets that include user credentials, crime metadata, risk matrices, and system logs. It is powered by MariaDB version 10 and is optimized for fast querying across both temporal and geographical dimensions. Access to this layer is tightly secured and managed through authenticated API calls with encrypted data transmissions, ensuring the confidentiality and integrity of sensitive information.

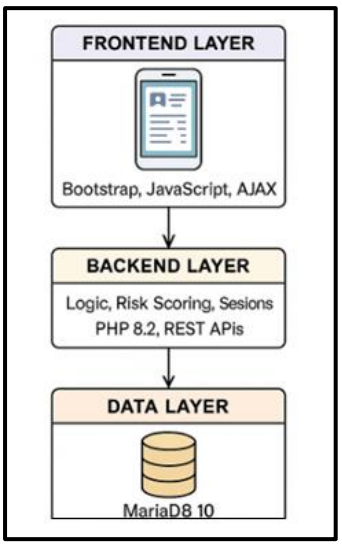


Figure 4.9 Conceptual Layered Architecture of ICRAS

4.7.2 System Design Workflow

The development of the ICRAS system adhered to a structured, phase based Software Development Lifecycle (SDLC), ensuring that each component of the system was systematically planned, developed, and deployed in accordance with both technical requirements and end user expectations. This approach was essential to meet the complex demands of real time crime risk analysis, which required not only precise data handling and robust processing capabilities but also user friendly interfaces and secure communication protocols. By adopting this lifecycle, the development team was able to break down the entire process into manageable phases, allowing for focused implementation, early detection of issues, and continuous improvement throughout the project. Each phase of the SDLC played a critical role in shaping the system's functionality and reliability. From the initial planning stage where key goals, system scope, and user needs were identified to the final deployment, every step was carefully documented and iteratively reviewed. This ensured that all architectural decisions were guided by practical insights from stakeholders, including field investigation officers, system testers, and IT administrators. The inclusion of iterative testing and refinement at each stage also contributed to higher software quality, reduced errors, and improved user satisfaction. In addition, the workflow emphasized modularity and scalability, allowing the system to accommodate future upgrades, feature expansion, and integration with external data sources or law enforcement platforms. Special consideration was also given to interoperability, responsiveness across mobile devices,

and secures access control to handle sensitive crime related data. Table 4.17 provides a summarized view of each development phase, along with the corresponding tools and technologies utilized to execute them efficiently, laying the foundation for a reliable and field ready mobile application.

Table 4.18
ICRAS System Development Phases and Tools

Development Phase	Description	Tools / Technology used
System Planning	Identification of system goals, integration needs, and user expectations	Stakeholder workshops, requirement matrix
Environment Setup	Installation of development tools, SDKs, and libraries	Android Studio, PHP 8.2, MariaDB 10, Swagger
Database & API Configuration	Schema design and establishment of secure communication endpoints	SQL, MariaDB, JSON, Swagger API
Frontend Development	Design of mobile interface and client-side interactions	Bootstrap 5, JavaScript, JQuery, AJAX, Mobiledetect.js
Backend Development	Implementation of core logic, user session control, and crime risk algorithms	PHP 8.2, REST API, MVC Architecture
App Compilation	Packaging of mobile application for testing and release	Android SDK, Gradle, APK/AAB generation
Deployment	Live deployment of backend services and app distribution to users	Private cloud server, Google Play Store

4.7.3 System Development

This flowchart visually represents the ICRAS Mobile Application Development Process in a structured and logical manner. It details the step-by-step approach, ensuring that all phases of the development lifecycle are systematically executed.

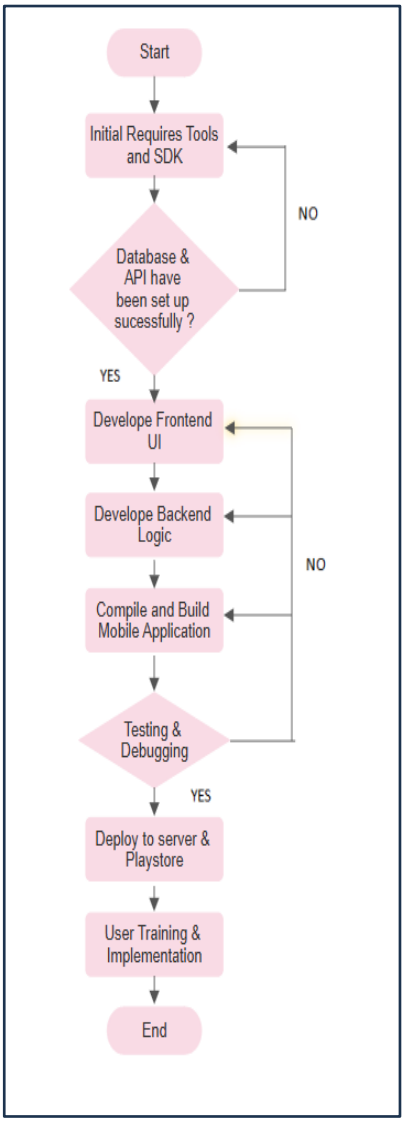


Figure 4.10 ICRAS Mobile Application Development Flowchart

The process begins with the Start phase, where project requirements, technical specifications, and system architecture are carefully planned. At this stage, developers define the key objectives, including the functionalities required for crime risk analysis, the integration of law enforcement databases, and the mobile app's usability for officers in the field. Once the project plan is established, the development team moves on to the installation of all required tools and SDKs, which forms the foundation for the entire

application. This step involves setting up essential programming environments, such as Android Studio for mobile development, PHP 8.2 for backend scripting, MariaDB 10 for database management, and Swagger API for secure communication. Additionally, Bootstrap 5, JQuery, JavaScript, AJAX, and Mobiledetect are configured to support a responsive and interactive user interface. This step ensures that all dependencies and software requirements are met before proceeding further. Following the installation, the process moves to a crucial decision point: Database & API setup. This phase verifies whether the database schema and API endpoints have been successfully configured. Since crime risk assessment relies heavily on structured data, it is essential to ensure that MariaDB 10 correctly stores and retrieves crime records, suspect details, and law enforcement actions. Furthermore, the Swagger API is implemented at this stage to ensure that communication between the mobile frontend and backend remains secure and efficient. If any misconfigurations occur, developers must revisit the database design and API settings to resolve errors. Once the database and API are properly configured, the next stage focuses on developing the frontend UI, which is a significant part of the user experience. Using Bootstrap 5 for styling, JQuery and JavaScript for interactivity, and AJAX for real time data updates, the frontend is designed to provide an intuitive and accessible interface for law enforcement officers. Additionally, Mobiledetect is incorporated to ensure that the app functions seamlessly across various devices, from smartphones to tablets. The frontend must be visually appealing while maintaining usability for officers who require quick access to crime data in real-time. After successfully designing the frontend, the process transitions into backend development, where the core logic of the application is built. Using PHP 8.2, developers write server-side scripts that handle crime data processing, suspect tracking, user authentication, and system operations. This step ensures that all API calls, database queries, and crime analytics functions are correctly executed. The backend also enables secure data exchange between officers, ensuring that sensitive crime information remains protected. If any errors arise in the backend, developers must revisit this step before proceeding to deployment. The next phase involves compiling and building the mobile application, where the entire system is assembled into a fully functional Android application. The developers generate an APK or AAB file, ensuring proper linkage between the UI, backend logic, database, and API services. This process ensures that the app is ready for deployment and testing. Once the application is compiled, it undergoes Testing & Debugging, a critical decision-making phase where the system is

evaluated for performance, security, and reliability. Developers perform unit testing to verify individual components, integration testing to ensure smooth interaction between different modules, and UI testing to validate the user experience. The mobile application is tested on multiple devices to check responsiveness, performance, and compatibility with different screen sizes. The crime data exchange processes are tested for efficiency and security to prevent data breaches. If any issues or bugs are found, the process loops back to either frontend, backend, or database setup, depending on the source of the problem. If the app passes all testing phases, it advances to deployment. In the deployment phase, the application is uploaded to a live server for API access and backend processing. Additionally, the mobile app is published on the Google Play Store or distributed internally to law enforcement agencies. This phase ensures that the application is accessible to officers and law enforcement personnel, enabling real-time crime risk assessment. Following deployment, the User Training & Implementation phase begins, where law enforcement officers receive hands-on training on how to effectively use the ICRAS mobile application. Training sessions focus on crime data reporting, suspect tracking, and real-time risk assessment, ensuring that users understand the app's functionalities. Officers test the app in live crime scenarios, and feedback is collected to make necessary improvements. Finally, the development process reaches the End phase, marking the successful completion of the mobile application deployment. At this stage, the ICRAS mobile app is fully operational and available for law enforcement agencies to use in real world crime risk assessment. However, future enhancements and updates may be required based on user feedback and evolving crime risk factors. Continuous monitoring and maintenance ensure that the application remains secure, up to date, and scalable for future crime prediction models. This structured development process ensures that the ICRAS mobile application is built with precision, allowing for seamless crime assessment, efficient law enforcement operations, and an improved crime prevention framework.

The ICRAS system was developed using a modular architecture that integrates ETL data processing, machine learning based risk scoring, and a user friendly interface for investigation officers. The development process involved iterative testing and refinement to ensure full functionality and operational reliability. Figure 4.11 presents the main ICRAS dashboard, which provides centralized access to crime data uploads, real time analytics, and predictive insights. This dashboard serves as the primary control

panel for officers to monitor incident trends, evaluate regional risk levels, and perform rapid assessments in the field.

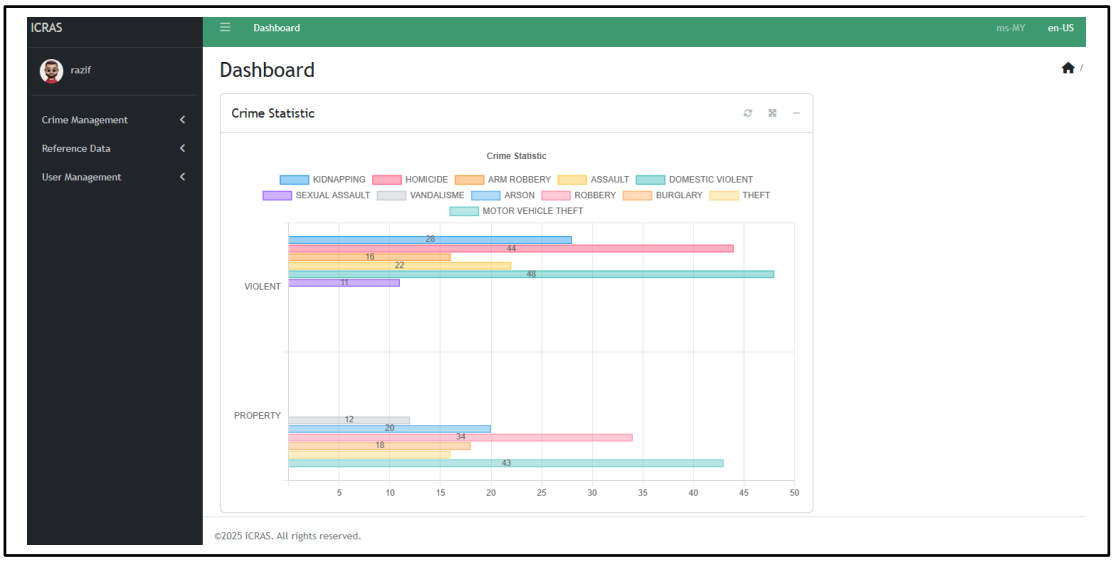


Figure 4.11 ICRAS Main Dashboard Interface

4.7.4 Authentication and Risk Assessment Flow User Process in ICRAS

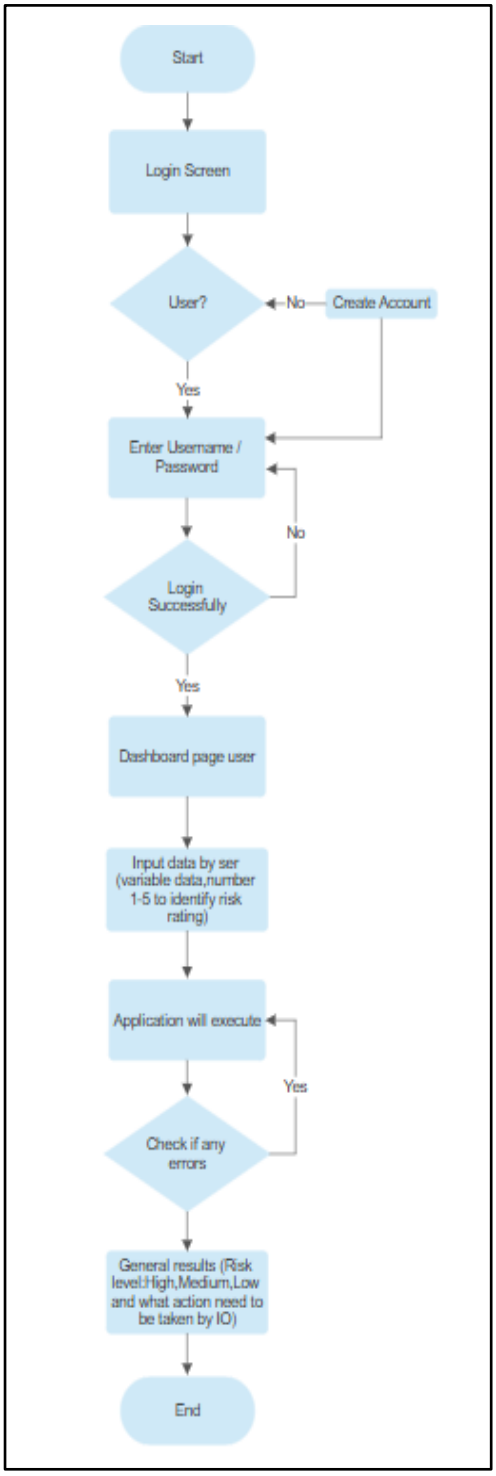


Figure 4.12 Flowchart of ICRAS Mobile Apps

The process begins at the Login Screen, where users initiate the system by either entering their credentials or creating a new account if they have not yet been registered. If the user is new, the system redirects them to the account creation module, where relevant identification details are entered for authentication approval. This step ensures that only authorized personnel gain access to the system, maintaining the integrity and confidentiality of all crime risk assessments. Authentication within ICRAS is implemented to guarantee that only verified law enforcement personnel can access sensitive operational data. *Figure 4.13* displays the login interface, where users authenticate using their assigned officer ID and secure password. This multi layered verification process safeguards all subsequent interactions within the system, preventing unauthorized access and ensuring that every user session is traceable and protected through secure communication protocols. Once an existing user enters their username and password, the system verifies the credentials against the database. If authentication fails, the process loops back, requiring users to re enter their details or retrieve their credentials through a recovery mechanism. Successful authentication grants access to the ICRAS Dashboard as illustrated in Figure 4.11, where officers can interact with crime data and input key variables relevant to risk evaluation. At this stage, the officer is required to input structured data, which includes a numeric risk rating (1–5) to quantify the severity and likelihood of a given crime incident. These variables serve as inputs for the system's risk assessment model, ensuring that predictions are based on measurable indicators rather than subjective interpretations. Once the data is entered, the application executes the predefined crime risk algorithm, which processes the inputted data and cross references it with historical crime trends, geographic factors, and severity parameters. This computational step is pivotal, as it transforms raw input into actionable intelligence. Following execution, the system undergoes an automated error-checking process, ensuring that inconsistencies or missing data points do not compromise the accuracy of risk predictions. If errors are detected, the process loops back, prompting the officer to correct the input before re executing the assessment. This iterative mechanism ensures data integrity and computational reliability, minimizing the risk of flawed evaluations. Upon successful execution, the final risk assessment report is generated as illustrated in Figure 4.14. The system classifies the risk level into three categories: High, Medium, or Low, enabling officers to make informed decisions on the appropriate course of action. Additionally, the system provides contextual guidance on necessary interventions, ensuring that law enforcement responses align

with the level of risk presented. By integrating real time data processing, automated risk evaluation, and use driven decision making, the ICRAS mobile application significantly enhances the precision and efficiency of crime risk management. This structured workflow ensures that crime risk assessments are not only data driven but also actionable, transparent, and adaptable to evolving security needs. Through a seamless combination of user authentication, structured data input, intelligent execution, and automated decision making, the system empowers law enforcement personnel with accurate, real time risk insights to enhance crime prevention and response strategies.

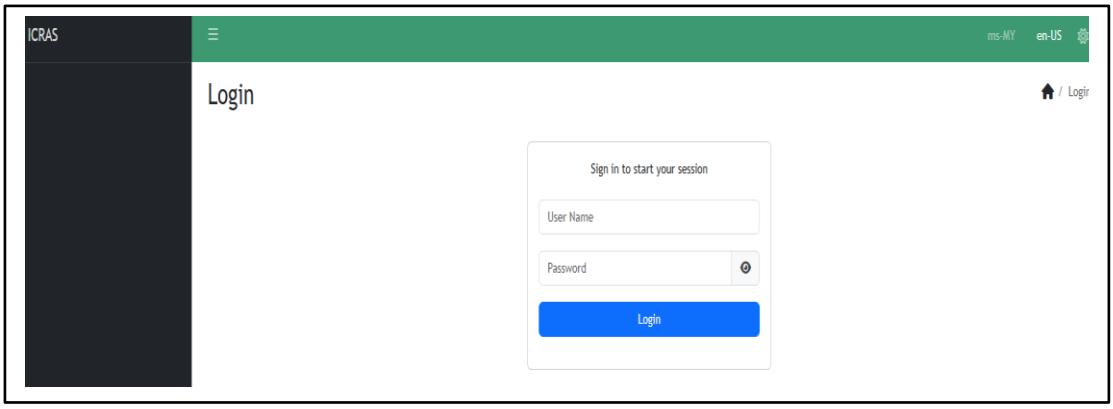


Figure 4.13 ICRAS Login Page

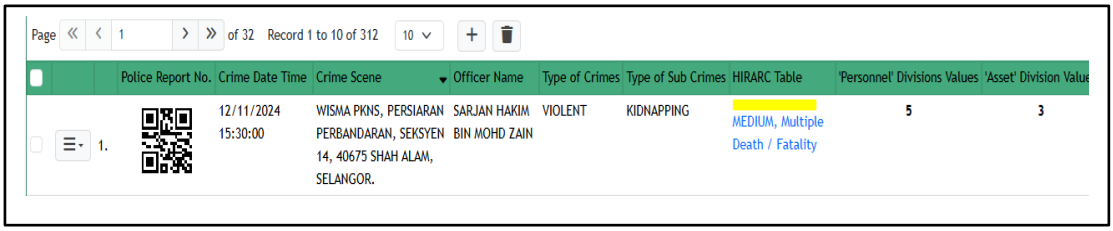


Figure 4.14 Risk Assessment and Scoring Interface

4.8 Implementation of the ICRAS Framework

4.8.1 Machine Learning Performance in Crime Prediction

The predictive component of the Integrated Crime Risk Assessment System (ICRAS) was implemented using the Random Forest algorithm to model and predict crime risk levels based on historical crime data and selected predictor variables. The objective of this modelling process was to evaluate the capability of the Random Forest method to learn patterns from the dataset and generate reliable crime risk predictions for operational use. The model was developed using the processed dataset obtained

through the ETL framework and trained to classify crime risk conditions based on the relationships between crime incidents and associated variables.

The performance of the Random Forest model is summarized in Table 4.18. The model achieved an accuracy of 92.5%, precision of 91.0%, recall of 90.2%, and an F1-score of 90.6%, indicating a high level of predictive reliability. These results demonstrate that the model successfully identified a high proportion of crime events while maintaining a balanced performance between correctly identifying crime-risk cases (precision) and detecting the majority of actual crime events (recall). The F1-score, which combines precision and recall into a single measure, further confirms that the Random Forest model achieved consistent and stable prediction performance.

Table 4.19
Machine Learning Results – Random Forest Framework

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	92.5	91.0	90.2	90.6

To ensure reliable evaluation of model performance, the dataset was divided using an 80:20 train-test split, where 80% of the data was used to train the Random Forest model, and the remaining 20% was reserved for testing. This approach enabled the predictive capability of the model to be assessed using previously unseen data. In addition, 5-fold cross-validation was conducted to evaluate the stability and generalizability of the model. The dataset was divided into five equal subsets, where the model was trained on four subsets and tested on the remaining subset. This procedure was repeated five times using different subsets for testing, ensuring that the performance results were not dependent on a single data partition.

A more detailed assessment of the prediction results is provided by the confusion matrix shown in Table 4.19. The confusion matrix summarizes the classification outcomes produced by the Random Forest model in terms of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). The Confusion Matrix for the Random Forest model is as follows:

Table 4.20
Confusion Matrix Table

	Predictive Positive	**Predictive Negative**
Actual Positive	TP = 1,084	FN = 208
Actual Negative	FP = 92	TN = 2,931

The confusion matrix shows that the model correctly identified 1,084 crime-risk cases (True Positives) and 2,931 non-crime-risk cases (True Negatives). A total of 208 crime-risk cases were not detected (False Negatives), while 92 non-crime cases were incorrectly classified as crime-risk cases (False Positives). These results indicate that the Random Forest model demonstrated strong capability in distinguishing between crime-risk and non-crime-risk conditions based on the input variables. Using the confusion matrix values, several performance metrics were calculated. Precision was computed as the ratio of True Positives to the sum of True Positives and False Positives, resulting in a precision value of 91.0%, indicating that when the model predicts a crime-risk event, it is correct approximately 91% of the time. Recall was calculated as the ratio of True Positives to the sum of True Positives and False Negatives, producing a recall value of 90.2%, which indicates that the model successfully detected the majority of actual crime-risk cases while missing approximately 9.8% of events. The F1-score, calculated as the harmonic mean of precision and recall, was 90.6%, demonstrating balanced performance between minimizing false positives and capturing true crime-risk events. The strong performance metrics, together with the k-fold cross-validation results, indicate that the Random Forest model is robust and generalizable across different partitions of the dataset, mitigating concerns of overfitting and demonstrating reliable predictive performance. The relatively low number of false positives and false negatives suggests that the selected predictor variables adequately capture crime risk conditions in the study area. Consequently, the model provides meaningful decision-support information for Investigation Officers when identifying cases or locations with elevated crime risk.

One of the key strengths of the Random Forest model is its robustness when applied to structured crime datasets. Crime data frequently contain noise, rare events, and irregular patterns that may negatively affect the performance of other prediction models. The ensemble structure of Random Forest reduces the influence of outliers and minimizes overfitting, allowing stable performance even when the data contain

inconsistencies. This characteristic makes the Random Forest model particularly suitable for operational crime prediction environments where data quality may vary. Despite the strong performance observed, several limitations should be acknowledged. The Random Forest algorithm is not specifically designed to capture sequential or temporal dependencies in crime data. Crime occurrences are often influenced by time-related factors such as seasonal patterns, day-of-week variations, and time-of-day effects. As a result, some temporal patterns may not be fully represented in the model predictions. Future improvements to the ICRAS framework may include the integration of temporal variables or the use of hybrid modelling approaches combining Random Forest with time-series models such as Recurrent Neural Networks (RNN) or Long Short-Term Memory (LSTM) networks. In addition, the Random Forest model demonstrated strong predictive performance with an accuracy of 92.5% and balanced precision and recall values. Its robustness, interpretability, and computational efficiency make it a suitable predictive modelling approach for the ICRAS framework. The results confirm that the Random Forest method can effectively transform historical crime data into reliable crime risk predictions that support evidence-based decision-making for Investigation Officers.

4.8.2 Pilot Testing of the ICRAS Framework

The researcher conducted a pilot test of the Integrated Crime Risk Assessment System (ICRAS) to evaluate how well the system functions in a real-world law enforcement setting. The test involved 30 participants, including law enforcement officers, crime analysts, and IT personnel from the Royal Malaysian Police. Over the course of two weeks, the participants were asked to use the Risk Appetite Scale to input crime-related data and then assess the resulting scores within the Crime Risk Matrix. The goal was to see how well the system helped users categorize crime risks as Low, Medium, or High based on input data. The pilot had several objectives: it aimed to assess whether the system effectively supported data-driven decision-making, whether the risk categories were clear and useful, and how well the system could integrate into the officers' and analysts' daily routines. Participants were asked to provide feedback on the usability of the system, the clarity of the risk assessments, and the overall functionality of the tools. Results from the pilot testing were largely positive. The response time for generating risk predictions averaged 2.4 seconds, a substantial

improvement over the previous manual process that took around 7.8 seconds. In terms of usability, participants rated the system highly, with an average score of 4.48 out of 5. This score suggests that the system was intuitive and easy for officers and analysts to use, even for those without extensive technical training. The Risk Appetite Index and Crime Risk Matrix were seen as practical tools for assessing crime risks, with participants reporting that the system enhanced their decision making by providing more structured and data driven insights. While feedback was overwhelmingly positive, a few participants suggested the addition of advanced filtering options to provide greater flexibility in the risk assessment process. However, these were minor recommendations and did not detract from the overall effectiveness of the system. The pilot confirmed that ICRAS was fit for real world application, validating its readiness for broader deployment across law enforcement agencies.

Table 4.21
Summary of Pilot Test Results for ICRAS

Evaluation Parameter	**Measurement**	**Result**	**Interpretation**
System Response Time	Time to generate risk predictions (seconds)	2.4 seconds	Significant reduction from the previous 7.8 seconds, demonstrating system efficiency.
Usability	Overall system usability (Likert scale 1–5)	4.48 / 5	High usability score, indicating the system is user friendly and easy to navigate.
Effectiveness of Risk Tools	Practicality of Risk Appetite Index & Crime Risk Matrix	Practical & Effective	Tools were deemed useful for assessing and categorizing crime risk in real time.
Qualitative Feedback	User comments on system functionality	Positive	Feedback highlighted the clarity of the system and suggested improvements, such as advanced filtering options.
Suitability	Fit of the system for daily workflows	High agreement (mean = 4.45)	Participants agreed that the system met their operational needs, improving decision-making.

4.9 Real Time Crime Data Access and Risk Analysis in ICRAS

4.9.1 System Usability and Performance Evaluation

The Integrated Crime Risk Assessment System (ICRAS) was piloted in Shah Alam, Selangor, as part of a broader initiative to modernize and optimize crime risk management. This deployment aimed to address key challenges faced by law enforcement, such as delays in crime data processing, low engagement among officers, and inaccuracies in traditional risk assessment methods. By integrating real time crime data streams from law enforcement databases and public sources, ICRAS provided a more comprehensive, responsive, and data driven approach to predictive policing. The objectives of the pilot implementation included reducing response time, enhancing user engagement, and improving risk assessment accuracy. The pilot involved system setup with integration of existing law enforcement databases, training sessions for officers, data collection for baseline and post implementation comparisons, and performance evaluations based on key metrics. The findings demonstrated significant improvements: response time was reduced from 7.8 seconds to 2.4 seconds, user engagement rate increased from 62% to 90%, and risk assessment accuracy improved from 78 % to 95%. These metrics, summarized in the table below, highlight the system's impact. The real-time data integration in ICRAS combines information from diverse sources such as law enforcement databases, social media, emergency calls, and community reports, offering a holistic view of crime trends. The user centric design of ICRAS, featuring an intuitive interface with customizable dashboards, real-time alerts, and significantly boosted officer engagement. Additionally, the use of advanced machine learning algorithms enabled the system to continuously adapt and improve its predictions based on new data inputs. Despite its success, the pilot highlighted challenges such as data quality issues, technical infrastructure requirements, and the need for comprehensive training. Addressing these challenges through data validation protocols, infrastructure upgrades, and scalable training programs will be crucial for future deployments. In conclusion, the pilot implementation of ICRAS in Shah Alam demonstrated its potential to revolutionize crime risk management in Malaysia by reducing response times, enhancing engagement, and improving risk assessment accuracy. This case study serves as a valuable example for adopting predictive policing technologies in developing

countries, providing lessons that can guide future implementations to maximize impact and contribute to a safer society.

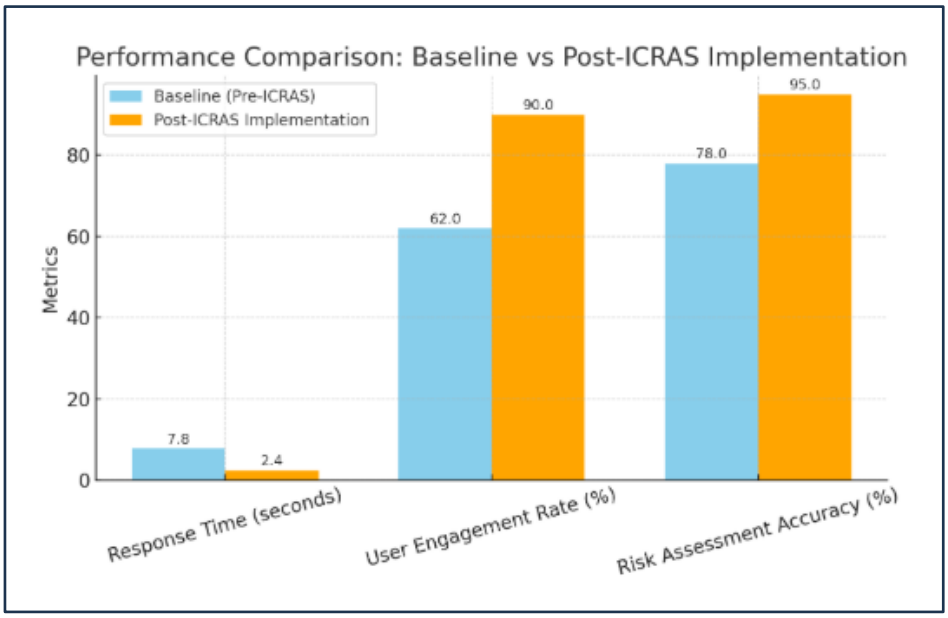


Figure 4.15 Performance Comparison Chart

4.9.2 Impact on Crime Prevention Strategies

The implementation of the Integrated Crime Risk Assessment System (ICRAS) demonstrates the practical value of data-driven approaches in supporting crime prevention strategies within the Royal Malaysian Police. The system provides Investigation Officers with real-time access to integrated crime data, predictive risk indicators, and structured risk classifications, enabling a more systematic approach to crime risk assessment. The results presented in Sections 4.5 and 4.8 show that the analytical components of ICRAS are capable of identifying meaningful crime patterns and classifying areas according to their level of crime risk. These capabilities support more informed decision-making and improve the effectiveness of operational planning. The quantitative findings indicate that crime risk levels are strongly associated with factors such as population density, unemployment rates, and police response times. By incorporating these variables into the Crime Risk Matrix and Risk Appetite Index, ICRAS enables Investigation Officers to identify areas with elevated risk levels and prioritize preventive actions accordingly. The predictive modelling results further demonstrate that historical crime data can be transformed into reliable indicators of future risk, allowing law enforcement personnel to anticipate potential crime

developments rather than relying solely on reactive responses. This shift from reactive investigation toward proactive crime prevention represents an important improvement in operational practice.

The availability of real-time data integration within ICRAS improves the ability of Investigation Officers to monitor crime trends and respond to emerging risks more efficiently. Instead of manually compiling information from multiple sources, officers can access consolidated crime data through a unified platform. This capability reduces delays in information processing and allows faster identification of high-risk situations. For example, when the system indicates an increase in property crime risk within a particular area, law enforcement personnel can respond by adjusting patrol deployment, strengthening surveillance activities, or implementing targeted prevention measures. Such data-driven responses improve resource allocation and increase the likelihood of preventing crime before it escalates.

The integration of predictive analytics within ICRAS also enhances the consistency of crime risk assessment. Traditional assessment methods often depend on individual experience and manual analysis, which may lead to variations in decision-making. By providing standardized risk scores and structured risk classifications, ICRAS supports more consistent evaluation of crime risk across different operational units. This consistency strengthens the reliability of decision-making and supports coordinated crime prevention strategies. The findings also demonstrate that the ICRAS platform can support long-term crime prevention planning. The ability to analyse historical patterns together with current data enables law enforcement agencies to identify persistent high-risk areas and recurring crime trends. These insights provide a basis for developing targeted prevention strategies and improving the strategic distribution of law enforcement resources. In this way, the system contributes not only to short-term operational decisions but also to longer-term planning and policy development.

While the implementation of ICRAS demonstrates significant benefits for crime prevention, several practical considerations remain important. The effectiveness of the system depends on the availability and quality of operational data, as well as the ability to integrate the platform with existing law enforcement information systems. In addition, appropriate training is required to ensure that Investigation Officers can effectively interpret and utilize the analytical outputs provided by the system. Addressing these factors is essential for ensuring the successful adoption and long-term

sustainability of the ICRAS framework. In addition, the results demonstrate that the ICRAS platform provides a practical and effective tool for supporting evidence-based crime prevention strategies. By integrating data processing, risk scoring, and predictive modelling into a unified system, ICRAS enables Investigation Officers to identify crime risks more systematically and respond more effectively to emerging threats. The findings confirm that the use of structured analytical tools can enhance operational efficiency and support proactive crime prevention within the Royal Malaysian Police.

4.9.3 Integration into Law Enforcement Practices

The integration of the Integrated Crime Risk Assessment System (ICRAS) into routine law enforcement practices represents an important step toward the adoption of data-driven crime risk assessment within the Royal Malaysian Police (RMP). The results presented in this study demonstrate that the ICRAS framework is capable of transforming crime data into structured risk information that can support operational decision-making. By providing Investigation Officers with access to integrated datasets, risk scores, and predictive indicators, the system supports a more systematic approach to crime investigation and resource deployment. The implementation of ICRAS enhances the daily workflow of Investigation Officers by reducing the need for manual data collection and analysis. Through the ETL-based data integration process, crime information from multiple sources can be consolidated into a unified platform, allowing officers to access relevant information more efficiently. This capability enables Investigation Officers to obtain a clearer overview of crime patterns and risk conditions within their operational areas, supporting faster and more consistent decision-making. The availability of structured risk classifications also assists officers in prioritizing investigations and allocating resources according to identified risk levels.

The analytical results presented in earlier sections demonstrate that crime risk can be reliably estimated using structured variables such as population density, unemployment rates, and historical crime records. The integration of these analytical outputs into operational practice allows Investigation Officers to incorporate quantitative risk indicators into their investigative processes. Instead of relying solely on individual experience or fragmented information sources, officers are able to support their decisions with consistent and data-based evidence. This approach improves both the transparency and reliability of crime risk assessments within operational

environments. The use of ICRAS within routine law enforcement practice also supports improved coordination among operational units. Since the system provides standardized risk information through a centralized platform, different units can access consistent data and share information more effectively. This capability facilitates coordinated responses to crime risks and reduces inconsistencies that may arise from separate data sources or individual assessment methods.

Successful integration of the ICRAS framework requires appropriate training and organizational support. Investigation Officers must be familiar with the interpretation of risk scores, predictive indicators, and system outputs in order to fully utilize the analytical capabilities of the platform. Training programs and operational guidelines are therefore important components in ensuring that the system can be effectively adopted in practice. In addition, continued system maintenance and periodic updates are necessary to ensure that the analytical models remain accurate and relevant as crime patterns evolve. The integration of ICRAS into law enforcement practices demonstrates how structured analytical tools can enhance investigative efficiency and support evidence-based policing. The system provides Investigation Officers with a practical mechanism for incorporating data-driven analysis into routine operations, improving the consistency and effectiveness of crime risk assessment within the Royal Malaysian Police. The findings indicate that the adoption of integrated analytical platforms such as ICRAS can strengthen operational decision-making and contribute to more effective crime prevention strategies.



Figure 4.16 Key Features of ICRAS for Investigation Officers

4.10 Concluding Remark

This study has made substantial contributions to both theoretical and practical domains, addressing critical challenges in crime risk assessment and leveraging advanced technological tools to enhance law enforcement strategies in Malaysia. The successful validation of the crime risk matrix approach within a Malaysian context, coupled with the integration of deep learning models into crime prediction, underscores the innovative nature of this research and its alignment with contemporary criminological advancements. The theoretical contributions are particularly significant as they address a gap in the existing literature, which has predominantly focused on Western urban settings. By applying and adapting the crime risk matrix to Malaysia, this research extends the findings of Weisburd et al. (2018), demonstrating that spatial and temporal dimensions of crime are equally relevant in non Western, rapidly urbanizing environments. This validation underscores the universality of these frameworks while contextualizing them for Southeast Asia, where unique socio-economic and cultural factors influence crime patterns. The findings not only validate the matrix's applicability but also open new avenues for exploring its integration into crime prevention strategies in similar developing countries. Equally significant is the integration of deep learning models into the development of the Integrated Crime Risk

Assessment System (ICRAS), which represents a forward-thinking contribution to the growing body of literature advocating for AI-driven policing. The work of Moses and Chan (2020) highlighted the transformative potential of artificial intelligence in augmenting traditional policing methods, and this study builds upon that foundation by demonstrating how machine learning algorithms can enhance crime prediction accuracy. The ability of ICRAS to process large datasets, identify patterns, and predict crime hotspots exemplifies the practical application of AI in criminology. This innovation not only advances theoretical understanding but also provides empirical evidence of the value of AI integration in real-world law enforcement contexts, bridging the gap between academic research and operational practice. The system's design ensures that it is not merely a technological tool but a comprehensive solution tailored to the needs of law enforcement officers, enabling them to move from reactive to proactive approaches in addressing crime risks. On the practical front, the development and implementation of ICRAS introduce a scalable and evidence-based model for crime risk assessment, fundamentally transforming the operational framework of law enforcement in Malaysia. Historically, crime prevention strategies have relied heavily on reactive policing methods, addressing incidents only after they have occurred. This approach, while effective in some instances, often overlooks underlying risk factors and fails to prevent future occurrences. ICRAS addresses these limitations by equipping the Royal Malaysia Police (RMP) with a proactive mechanism to assess and mitigate risks before crimes escalate. As supported by [230], evidence based policing strategies are crucial in modern law enforcement, and this research provides a practical demonstration of how data analytics can inform decision making. By analyzing crime patterns, predicting high-risk areas, and enabling strategic resource allocation, ICRAS positions law enforcement to anticipate threats and implement timely interventions. The system’s ability to provide real-time crime data analytics ensures that officers can respond swiftly to emerging threats, a capability that is increasingly critical in the context of Malaysia’s rapid urbanization and population growth.

CHAPTER 5
CONCLUSION

5.1 Summary of Key Findings

This study addressed three main objectives: (1) the development of a Risk Appetite Index and Crime Risk Matrix for structured crime risk evaluation, (2) the design of the Integrated Crime Risk Assessment System (ICRAS) framework, and (3) the implementation of a functional ICRAS platform to support Investigation Officers (IOs) within the Royal Malaysian Police (RMP). The findings demonstrate that the integration of structured risk assessment methods, data analytics, and predictive modelling can significantly improve the consistency and effectiveness of crime risk evaluation. The first objective focused on establishing a structured approach for assessing crime risk through the development of the Risk Appetite Index and Crime Risk Matrix. Quantitative analysis using logistic regression and decision tree techniques demonstrated that crime probability is strongly associated with key variables such as population density, unemployment rates, and police response times. These results confirmed that the selected variables provide meaningful indicators of crime risk and support the use of structured risk scoring for identifying high-risk conditions. The Risk Appetite Index and Crime Risk Matrix therefore provide a practical framework for classifying crime risk levels and supporting systematic evaluation of crime exposure within operational environments.

The second objective involved the design and development of the ICRAS framework as an integrated analytical platform for crime risk assessment. The framework incorporates an ETL-based data integration process that consolidates information from multiple sources into a unified database structure. This integration enables consistent data processing and supports the analytical components of the system. Predictive modelling using the Random Forest algorithm demonstrated strong performance in classifying crime risk levels, confirming the suitability of the modelling approach for moderate-sized and heterogeneous crime datasets. The analytical results indicate that historical crime data can be effectively transformed into structured risk indicators that support evidence-based decision-making.

The third objective focused on the development and implementation of the ICRAS platform for operational use by Investigation Officers. The resulting system provides access to integrated crime data, structured risk scores, and visualization tools that support the interpretation of crime risk conditions. The platform enables Investigation Officers to identify high-risk areas and emerging patterns more efficiently, supporting improved prioritization of investigations and resource allocation. The findings demonstrate that the integration of predictive modelling and structured risk evaluation within a single platform enhances the ability of law enforcement personnel to conduct systematic and data-driven crime risk assessments. The research adopted a mixed-methods approach that combined quantitative modelling with qualitative user feedback to ensure both analytical validity and operational relevance. Quantitative methods were used to validate the risk scoring model and predictive components, while qualitative feedback from Investigation Officers provided insights into system usability and practical implementation. The ETL process played an important role in ensuring data consistency and reliability, supporting accurate analysis and model development.

The pilot implementation of ICRAS within the Royal Malaysian Police demonstrated the practical applicability of the system in operational environments. Feedback from Investigation Officers indicated that the platform improved accessibility to crime information and supported more efficient analytical processes (Appendix 1). The positive evaluation from users suggests that the system provides meaningful support for daily investigative activities and has potential for broader adoption within law enforcement operations. The findings of this study demonstrate that the integration of structured risk assessment methods, data integration techniques, and predictive analytics within the ICRAS framework provides a reliable and scalable approach to crime risk assessment. The research establishes a practical foundation for the adoption of data-driven crime risk evaluation within the Royal Malaysian Police and supports the development of more systematic and proactive crime prevention strategies.

5.2 Contributions to Knowledge and Practice

This research makes significant contributions to both theoretical and practical domains by introducing a novel, data-driven crime risk assessment framework tailored specifically for the Royal Malaysian Police (RMP). The primary theoretical

contribution lies in the development and validation of the Risk Appetite Index and Crime Risk Matrix, which advance existing criminological models by integrating real-time data analytics, spatial-temporal forecasting, and AI-driven predictions. These innovative methodologies not only enhance traditional crime risk assessment frameworks but also extend their applicability beyond Western-centric models, particularly in the context of Malaysian law enforcement. This work contextualizes predictive policing methodologies within a Malaysian framework, ensuring the relevance of crime prediction models in non-Western settings.

From a practical standpoint, the implementation of ICRAS represents a major advancement in crime risk evaluation, offering a scalable and adaptive model for proactive crime prevention. By systematically integrating law enforcement databases, geospatial crime mapping, and machine learning algorithms, ICRAS transforms conventional crime response approaches into more dynamic, data-driven crime prevention strategies. This new approach facilitates real-time crime predictions, enabling law enforcement agencies to shift from reactive to proactive policing.

The research also underscores the potential of automated risk scoring models to enhance law enforcement decision-making by optimizing patrol scheduling, improving resource allocation, and strengthening interdepartmental collaboration. By providing law enforcement agencies with timely, actionable insights, ICRAS aids in making informed decisions that are crucial in responding to and mitigating crime before it escalates.

Additionally, the study has policy-level implications for national crime prevention initiatives in Malaysia. By equipping law enforcement with real-time risk insights and predictive capabilities, ICRAS supports the transition toward a data-driven, technology-enabled law enforcement model. This aligns with global trends in smart policing and AI-driven crime prevention, positioning Malaysia at the forefront of innovative law enforcement methodologies that could potentially be adopted across other jurisdictions.

5.3 Recommendations for Future Research

Although this study has made significant progress in advancing crime risk prediction through the development of the Integrated Crime Risk Assessment System (ICRAS), there remain substantial opportunities for further enhancement and broader

application. The findings of this research provide several important areas for future work, each of which could enhance the system's predictive capabilities and extend its relevance beyond the scope of this study. One of the most critical areas for improvement is the integration of temporal dynamics into the ICRAS framework. While the Random Forest model used in this study has proven effective in static prediction tasks, its capacity to capture crime trends over time is limited. Future research could explore the integration of more advanced time-series models, such as Recurrent Neural Networks (RNNs) or Long Short-Term Memory (LSTM) networks, which are known for their ability to learn from time-dependent patterns. By incorporating these techniques, ICRAS could more effectively identify emerging crime trends, predict their peaks and declines, and, most importantly, support earlier interventions and improved resource allocation. This would elevate the system's predictive power, especially in environments where crime exhibits cyclical or seasonal patterns.

Moreover, expanding the range of data sources used by ICRAS is another important avenue for future research. While the current version relies heavily on structured police data and socio-economic indicators, incorporating additional data such as local event information, transportation patterns, and social media trends could provide a more comprehensive view of the factors influencing crime. The inclusion of spatial data, such as satellite imagery or GPS-based movement patterns, would also enhance the system's accuracy, particularly in terms of hotspot detection and situational awareness at the local level. Integrating diverse data streams would ensure that ICRAS remains a robust tool for dynamic crime forecasting and more accurately reflects the complexity of real-world crime patterns. Testing and adapting the ICRAS framework in different socio-economic and cultural contexts also presents a valuable research direction. While this study focused on the Royal Malaysian Police (RMP), applying ICRAS in diverse settings, such as urban and rural areas or different countries, would provide valuable insights into its flexibility and generalizability. Comparing ICRAS's performance across various environments would not only test the robustness of the model but also highlight which contextual factors most significantly influence its outcomes. This could include exploring how the model adapts to varying local crime trends, demographic factors, and law enforcement practices.

Furthermore, user experience remains a crucial aspect of ICRAS's success. While Investigation Officers (IOs) participated in the initial pilot phase, further research is needed to assess the usability and adoption of ICRAS in real-world settings. Future

studies should investigate how the system can be better integrated into the officers' daily routines, improving its intuitiveness through design enhancements such as simplified navigation, voice-assisted inputs, and interactive visual dashboards. Understanding and addressing the cultural and institutional barriers to technology acceptance within law enforcement agencies will be vital to ensure the widespread adoption and effective use of ICRAS. Another promising direction for future research involves the integration of ICRAS with other operational systems used by the RMP and similar agencies. Linking ICRAS with decision-support tools for case management, resource allocation, and emergency response would create a more interconnected and intelligent policing environment. However, this integration would necessitate addressing technical and organizational challenges, including data compatibility, system security, and workflow integration. Ensuring smooth integration across disparate systems will be key to maximizing ICRAS's effectiveness and operational relevance.

As machine learning tools become increasingly embedded in law enforcement practices, the ethical implications of predictive policing systems must be thoroughly examined. Future research should focus on mitigating algorithmic bias, ensuring transparency in system evaluations, and establishing ethical guidelines for the responsible use of predictive models in policing. Addressing these ethical concerns is essential not only for improving the operational efficiency of ICRAS but also for ensuring that it upholds fairness and accountability in law enforcement practices. Efforts should be made to ensure that ICRAS is developed and deployed in a manner that respects privacy and human rights, especially in sensitive crime prediction scenarios.

This study has laid an important foundation for data-driven crime risk assessment, but much work remains to be done. Enhancing ICRAS's temporal modeling capabilities, diversifying its data sources, deepening user engagement, improving system interoperability, and addressing the ethical considerations surrounding predictive policing will be crucial in refining ICRAS and maximizing its real-world impact. Continued research in these areas will help shape the future of intelligent, evidence-based policing, enhancing the role of technology in advancing public safety and enabling law enforcement agencies to proactively manage crime risks. In addition to these technical and operational areas, it is also critical to address the organizational and policy challenges that currently limit the official adoption of ICRAS within the Royal Malaysian Police (RMP). A specific recommendation for future action is the need

for comprehensive policy reforms that would facilitate the integration of predictive policing systems like ICRAS into the RMP's standard operating procedures. Such reforms should establish a formal policy framework that not only supports the adoption of ICRAS but also ensures its compliance with ethical standards, data privacy laws, and transparency in its implementation. Furthermore, the policy should outline clear protocols for training officers, integrating the system into daily operations, and scaling the system across various jurisdictions within the RMP. By adopting these policy changes, the RMP would be better positioned to harness the full potential of ICRAS, using it to proactively manage crime risks, optimize resource allocation, and ultimately improve public safety outcomes.

REFERENCES

[1] D. T. L. Shek, "Protests in Hong Kong (2019–2020): A perspective based on quality of life and well-being," *Applied Research in Quality of Life*, vol. 15, no. 3, pp. 619–635, 2020. doi: 10.1007/s11482-020-09825-2.

[2] J. Zhao, C. Zhu, and S. Zhao, "Comparing the spatiotemporal dynamics of urbanization in moderately developed Chinese cities over the past three decades: Case of Nanjing and Xi'an," *Journal of Urban Planning and Development*, vol. 141, no. 4, 2015. doi: 10.1061/(ASCE)UP.1943-5444.0000251.

[3] M. Felson, S. Jiang, and Y. Xu, "Routine activity effects of the COVID-19 pandemic on burglary in Detroit, March 2020," *Crime Science*, vol. 9, no. 1, 2020. doi: 10.1186/s40163-020-00120-x.

[4] A. Hasenclever, P. Mayer, and V. Rittberger, "Integrating theories of international regimes," *Review of International Studies*, vol. 26, no. 1, pp. 3–33, 2000. doi: 10.1017/S0260210500000036.

[5] M. Baker and N. Westelius, "Crime, expectations, and the deterrence hypothesis," in *Handbook on the Economics of Crime*, 2013. doi: 10.4337/9781781000151.00017.

[6] J. Graan, "A case study of non-violent property crime victimisation in a South African urban residential neighbourhood: Exploring the excessive use of force and destruction caused by burglars to gain entry to victims' properties," *International Journal of Criminology and Sociology*, vol. 10, pp. 992–1005, 2021. doi: 10.6000/1929-4409.2021.10.117.

[7] D. Gingerich and V. Oliveros, "Police violence and the underreporting of crime," *Economics and Politics*, vol. 30, no. 1, pp. 78–105, 2017. doi: 10.1111/ecpo.12102.

[8] S. Brand and R. Price, *The Economic and Social Costs of Crime*. London, U.K.: Home Office Research Study, 2000.

[9] B. Perry and S. Alvi, "'We are all vulnerable': The in terrorem effects of hate crimes," *International Review of Victimology*, vol. 18, no. 1, pp. 57–71, 2012. doi: 10.1177/0269758011422475.

[10] T. D. Crowe and L. J. Fennelly, "Crime and loss prevention," in *Crime Prevention Through Environmental Design*, 3rd ed. Oxford, U.K.: Butterworth-Heinemann, 2013, pp. 105–125. doi: 10.1016/B978-0-12-411635-1.00007-3.

[11] W. Ye-Jun, “Strengthening intelligence education with information-processing and knowledge-organization competencies,” *Journal of Strategic Security*, vol. 6, no. 3, pp. 10–24, 2013. doi: 10.5038/1944-0472.6.3.2.

[12] M. A. Lasi and E. B. Yunusi, “Factors influencing high crime rate in Kuala Lumpur urban community,” *International Journal of Publication and Social Studies*, vol. 5, no. 2, pp. 147–153, 2020. doi: 10.18488/journal.135.2020.52.147.153.

[13] M. C. Soh, “Crime and urbanization: Revisited Malaysian case,” *Procedia - Social and Behavioral Sciences*, vol. 42, pp. 291–299, 2012. doi: 10.1016/j.sbspro.2012.04.193.

[14] Z. Islam, “Dynamics of rural crime in South Asia: A criminological overview,” *IJSSTEM*, vol. 1, no. 2, 2024. doi: 10.59781/njyn3048.

[15] Royal Malaysian Police, *Annual Crime Statistics Report 2019*. Bukit Aman, Malaysia: Royal Malaysian Police, 2019.

[16] L. K. Cheng and Y. K. Kwan, “What are the determinants of the location of foreign direct investment? The Chinese experience,” *Journal of International Economics*, vol. 51, no. 2, pp. 379–400, 2000. doi: 10.1016/S0022-1996(99)00032-X.

[17] H. Matakovic and I. Cunjak Matakovic, “The impact of crime on security in tourism,” *Security and Defence Quarterly*, vol. 27, no. 5, pp. 1–20, 2019.

[18] M. I. Mud Shukri *et al.*, “Prevalence and factors associated with depression, anxiety and stress in Malaysia during COVID-19 pandemic: A systematic review,” *PLoS One*, vol. 18, no. 7, 2023. doi: 10.1371/journal.pone.0288618.

[19] J. H. Ratcliffe, “How the police can impact crime,” in *Reducing Crime*. London, U.K.: Routledge, 2018, pp. 101–125. doi: 10.4324/9781351132350-6.

[20] Royal Malaysian Police, *Organizational Structure of the Royal Malaysian Police*. Bukit Aman, Malaysia: Royal Malaysian Police, 2023.

[21] Ministry of Home Affairs Malaysia, *Law Enforcement and Crime Management Framework*. Putrajaya, Malaysia: Ministry of Home Affairs, 2022.

[22] R. Weitzer, “Incidents of police misconduct and public opinion,” *Journal of Criminal Justice*, vol. 30, no. 5, pp. 397–408, 2002. doi: 10.1016/S0047-2352(02)00150-2.

[23] R. Orr *et al.*, “Police fitness: An international perspective on current and future challenges,” *Sports*, vol. 13, no. 7, 2025. doi: 10.3390/sports13070219.

[24] M. Maidin, M. Hussin, N. Zakaria, S. Rahim, and K. Samsudin, “The need for integrated crime risk assessment systems using data-driven strategies for crime prevention in Malaysia,” *International Journal of Academic Research in Business and Social Sciences*, vol. 14, no. 11, 2024. doi: 10.6007/IJARBSS/v14-i11/23600.

[25] “Predictive policing: The role of AI in crime prevention,” *International Journal of Computer Applications Technology and Research*, 2024. doi: 10.7753/IJCATR1310.1006.

[26] R. Marasambessy, “Information technology-based law enforcement in increasing public trust in the police,” *Gema Wiralodra*, vol. 14, no. 2, pp. 790–798, 2023. doi: 10.31943/gw.v14i2.507.

[27] M. Ali, B. Mishra, D. Thakker, S. Mazumdar, and S. Simpson, “Using citizen science to complement IoT data collection: A survey of motivational and engagement factors in technology-centric citizen science projects,” *IoT*, vol. 2, no. 2, pp. 275–309, 2021. doi: 10.3390/iot2020015.

[28] A. H. Alkon, “Place, stories, and consequences,” *Organization & Environment*, vol. 17, no. 2, pp. 145–169, 2004. doi: 10.1177/1086026604264881.

[29] W. D. and W. N., “Community-oriented policing: A theoretical exploration and its implications for building safer communities,” *International Journal of Research and Innovation in Social Science*, vol. VIII, no. II, pp. 15–21, 2024. doi: 10.47772/IJRISS.2024.802002.

[30] M. Hanslmaier, “Crime, fear and subjective well-being: How victimization and street crime affect fear and life satisfaction,” *European Journal of Criminology*, vol. 10, no. 5, pp. 515–533, 2013. doi: 10.1177/1477370812474545.

[31] A. Razip *et al.*, “A mobile visual analytics approach for law enforcement situation awareness,” in *Proc. IEEE Pacific Visualization Symposium*, 2014, pp. 169–176. doi: 10.1109/PacificVis.2014.54.

[32] R. Mission, “Spatial analysis leveraging geographic information system and kernel density estimation for crime prediction,” *JITC*, vol. 6, no. 1, pp. 95–104, 2024. doi: 10.69478/jitc2024v6n2a10.

[33] I. Mbewu, E. E. Obioha, and I. Mugari, “Encouraging and discouraging factors to crime reporting in South Africa: A survey of residents' crime reporting behaviour in Mthatha South Africa Police Service precinct,” *Cogent Social Sciences*, vol. 7, no. 1, 2021. doi: 10.1080/23311886.2021.2002542.

[34] J. Liu, “Enhancing English language education through big data analytics and generative AI,” *Journal of Web Engineering*, pp. 227–250, 2024. doi: 10.13052/JWE1540-9589.2322.

[35] S. Gle, E. Silverman, and T. Weigend, “If robots cause harm, who is to blame? Self-driving cars and criminal liability,” *New Criminal Law Review*, vol. 19, no. 3, pp. 412–436, 2016. doi: 10.1525/NCLR.2016.19.3.412.

[36] R. Simmons, “Quantifying criminal procedure: How to unlock the potential of big data in our criminal justice system,” *SSRN Electronic Journal*, 2016. doi: 10.2139/ssrn.2816006.

[37] J. A. Singer, M. K. Miller, and M. Adya, “The impact of DNA and other technology on the criminal justice system: Improvements and complications,” *Albany Law Journal of Science & Technology*, vol. 17, pp. 87–114, 2007.

[38] H. Pieterse, M. Olivier, and R. Heerden, “Evaluation framework for detecting manipulated smartphone data,” *SAIEE Africa Research Journal*, vol. 110, no. 2, pp. 67–76, 2019. doi: 10.23919/SAIEE.2019.8732797.

[39] M. Dugato, “Understanding the multi-level interactions between physical environment and neighborhood characteristics in assessing vulnerability to crime in micro-places,” *Crime & Delinquency*, vol. 68, no. 11, pp. 1923–1947, 2021. doi: 10.1177/00111287211035984.

[40] M. D. Reisig, “Community and problem-oriented policing,” *Crime and Justice*, vol. 39, no. 1, pp. 1–53, 2010. doi: 10.1086/652384.

[41] A. Farrell, V. Bouch, and D. Wolfe, “Assessing the impact of state human trafficking legislation on criminal justice system outcomes,” *Law & Policy*, vol. 41, no. 2, pp. 174–197, 2019. doi: 10.1111/lapo.12124.

[42] L. Ennis, T. Hargreaves, and M. Gulayets, “The integrated threat and risk assessment centre: A program evaluation investigating the implementation of threat management recommendations,” *Journal of Threat Assessment and Management*, vol. 2, no. 2, pp. 114–126, 2015. doi: 10.1037/tam0000040.

[43] M. Hamilton, “Judicial gatekeeping on scientific validity with risk assessment tools,” *Behavioral Sciences & the Law*, vol. 38, no. 3, pp. 226–245, 2020. doi: 10.1002/bsl.2456.

[44] S. Goel, R. Shroff, J. Skeem, and C. Slobogin, “The accuracy, equity, and jurisprudence of criminal risk assessment,” in *Handbook of Risk Assessment in Criminal Justice*, 2021. doi: 10.4337/9781788972826.00007.

[45] A. Day, E. Darwinkel, and J. Vess, “The characteristics of registered sexual offenders in an Australian jurisdiction,” *Journal of Investigative Psychology and Offender Profiling*, vol. 14, no. 2, pp. 120–132, 2016. doi: 10.1002/jip.1461.

[46] J. Looman and J. Abracen, “The risk need responsivity model of offender rehabilitation: Is there really a need for a paradigm shift?,” *International Journal of Behavioral Consultation and Therapy*, vol. 8, no. 3–4, pp. 30–36, 2013. doi: 10.1037/h0100980.

[47] J. Coldren, A. Huntoon, and M. Medaris, “Introducing smart policing: Foundations, principles, and practice,” *Police Quarterly*, vol. 16, no. 3, pp. 275–286, 2013. doi: 10.1177/1098611113497042.

[48] M. Oswald, J. Grace, S. Urwin, and G. C. Barnes, “Algorithmic risk assessment policing models: Lessons from the Durham HART model and ‘experimental’ proportionality,” *Information & Communications Technology Law*, vol. 27, no. 2, pp. 223–250, 2018. doi: 10.1080/13600834.2018.1458455.

[49] J. Grace, “‘AI theory of justice’: Using Rawlsian approaches to better legislate on machine learning in government,” *SSRN Electronic Journal*, 2020. doi: 10.2139/ssrn.3588256.

[50] R. D. Hoge and D. A. Andrews, *Youth Level of Service/Case Management Inventory 2.0*, PsycTESTS Dataset, 2012. doi: 10.1037/t05078-000.

[51] R. Borum, H. P. Lodewijks, P. Bartel, and A. E. Forth, “The structured assessment of violence risk in youth (SAVRY),” in *Handbook of Violence Risk Assessment*, 2020, pp. 438–461. doi: 10.4324/9781315518374-22.

[52] T. Karppi, “‘The computer said so’: On the ethics, effectiveness, and cultural techniques of predictive policing,” *Social Media + Society*, vol. 4, no. 2, 2018. doi: 10.1177/2056305118768296.

[53] D. Birks, M. Townsley, and T. Hart, “Predictive policing in an Australian context: Assessing viability and utility,” 2023. doi: 10.52922/ti78870.

[54] E. Maeder and R. Wiener, “Narrowly tailored actuarial models for affirmative action in higher education,” *Analyses of Social Issues and Public Policy*, vol. 10, no. 1, pp. 116–145, 2010. doi: 10.1111/j.1530-2415.2010.01219.

[55] M. Islam, “Exploring ethical dimensions in AI: Navigating bias and fairness in the field,” *JAIGS*, vol. 1, no. 1, pp. 13–17, 2024. doi: 10.60087/jaigs.v1i1.p18.

[56] J. P. Singh and S. Fazel, “Forensic risk assessment,” *Criminal Justice and Behavior*, vol. 37, no. 9, pp. 965–988, 2010. doi: 10.1177/0093854810374274.

[57] D. Lovallo and O. Sibony, “Broadening the frame: How behavioral strategy redefines strategic decisions,” *Strategy Science*, vol. 3, no. 4, pp. 658–667, 2018. doi: 10.1287/stsc.2018.0071.

[58] L. Pedersen, K. Rasmussen, and P. Elsass, “Risk assessment: The value of structured professional judgments,” *International Journal of Forensic Mental Health*, vol. 9, no. 2, pp. 74–81, 2010. doi: 10.1080/14999013.2010.499556.

[59] P. Skrivankov *et al.*, “Violence risk and personality assessment in adolescents by Structured Assessment of Violence Risk in Youth (SAVRY) and High School Personality Questionnaire (HSPQ): Focus on protective factors strengthening,” *Frontiers in Psychiatry*, vol. 13, 2023. doi: 10.3389/fpsyt.2022.1067450.

[60] C. Patel *et al.*, “Strategies for redesigning withdrawn drugs to enhance therapeutic efficacy and safety: A review,” *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 15, no. 1, 2025. doi: 10.1002/wcms.70004.

[61] N. Duits, S. van der Hoorn, M. Wiznitzer, R. M. Wettstein, and E. de Beurs, “Quality improvement of forensic mental health evaluations and reports of youth in the Netherlands,” *International Journal of Law and Psychiatry*, vol. 35, no. 5–6, pp. 440–444, 2012. doi: 10.1016/j.ijlp.2012.09.018.

[62] N. Duits, T. Doreleijers, and W. van den Brink, “Assessment of violence risk in youth for juvenile court: Relevant factors for clinical judgment,” *International Journal of Law and Psychiatry*, vol. 31, no. 3, pp. 236–240, 2008. doi: 10.1016/j.ijlp.2008.04.007.

[63] L. Sanchez-Pinto and T. Bennett, “Evaluation of machine learning models for clinical prediction problems,” *Pediatric Critical Care Medicine*, vol. 23, no. 5, pp. 405–408, 2022. doi: 10.1097/PCC.0000000000002942.

[64] J. Rane, K. Kaya, S. Mallick, and N. Rane, “Scalable and adaptive deep learning algorithms for large-scale machine learning systems,” 2024. doi: 10.70593/978-81-981271-0-5_2.

[65] F. Khan, B. Das, and P. Mohammad, “Urban growth modeling and prediction of land use land cover change over Nagpur city, India using cellular automata approach,” in *Advances in Geographical and Environmental Sciences*, 2022, pp. 261–282. doi: 10.1007/978-981-16-7373-3_13.

[66] A. Joshi, A. Singh, S. Chauhan, and A. Sharma, “Decision tree algorithm for credit card fraud detection,” *World Wide Web*, vol. 18, no. 4, 2021. doi: 10.29121/web/v18i4/103.

[67] X. Su-Min, L. Liu, G. Sun, B. Pan, L. Lin, and P. Guo, “Enhanced e-commerce fraud prediction based on a convolutional neural network model,” *Computers, Materials & Continua*, vol. 75, no. 1, pp. 1107–1117, 2023. doi: 10.32604/cmc.2023.034917.

[68] D. Ganatra, “The sensitivity conundrum — Random forest or boosting,” *International Journal of Emerging Trends in Engineering Research*, vol. 8, no. 7, pp. 3223–3227, 2020. doi: 10.30534/IJETER/2020/56872020.

[69] Y. Liang, C. Chen, T. Tian, and K. Shu, “Fair classification via domain adaptation: A dual adversarial learning approach,” *Frontiers in Big Data*, vol. 5, 2023. doi: 10.3389/fdata.2022.1049565.

[70] “Utilizing predictive insights for future planning: Redefining choices with advanced data solutions,” *International Journal of Computer Applications Technology and Research*, 2025. doi: 10.7753/IJCATR1402.1004.

[71] R. Galanti, B. Coma-Puig, M. de Leoni, J. Carmona, and N. Navarin, “Explainable predictive process monitoring,” 2020. doi: 10.48550/arXiv.2008.01807.

[72] W. Rizzi, C. Di Francescomarino, and F. M. Maggi, “Explainability in predictive process monitoring: When understanding helps improving,” in *Lecture Notes in Business Information Processing*, 2020, pp. 141–158. doi: 10.1007/978-3-030-58638-6_9.

[73] R. Sindhgatta, C. Ouyang, and C. Moreira, “Exploring interpretability for predictive process analytics,” in *Lecture Notes in Computer Science*, 2020, pp. 439–447. doi: 10.1007/978-3-030-65310-1_31.

[74] H. Kiavarz, M. Jadidi, A. Rajabifard, and G. Sohn, “BIM-GIS oriented intelligent knowledge discovery,” *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-4/W10, pp. 79–82, 2018. doi: 10.5194/isprs-archives-xlii-4-w10-79-2018.

[75] A. Tounsi and M. Temimi, “A systematic review of natural language processing applications for hydrometeorological hazards assessment,” *Natural Hazards*, vol. 116, no. 3, pp. 2819–2870, 2023. doi: 10.1007/s11069-023-05842-0.

[76] V. Rotaru, Y. Huang, T. Li, J. Evans, and I. Chattopadhyay, “Event-level prediction of urban crime reveals a signature of enforcement bias in U.S. cities,” *Nature Human Behaviour*, vol. 6, no. 8, pp. 1056–1068, 2022. doi: 10.1038/s41562-022-01372-0.

[77] L. Duan, “Performance evaluation and practical use of supervised data mining algorithms for credit card approval,” in *Proc. 2020 International Conference on Computing and Data Science (CDS)*, 2020, pp. 251–254. doi: 10.1109/CDS49703.2020.00057.

[78] Y. Zhou, M. Kantarcioglu, and C. Clifton, “Improving fairness of AI systems with lossless de-biasing,” 2021. doi: 10.48550/arXiv.2105.04534.

[79] W. C. McGaghie and C. D. Kreiter, “Special article: Holistic versus actuarial student selection,” *Teaching and Learning in Medicine*, vol. 17, no. 1, pp. 89–91, 2005. doi: 10.1207/s15328015tlm1701_16.

[80] J. McLaughlin *et al.*, “Quantitative assessment of the heterogeneity of PD-L1 expression in non-small-cell lung cancer,” *JAMA Oncology*, vol. 2, no. 1, p. 46, 2016. doi: 10.1001/jamaoncol.2015.3638.

[81] M. D. Cavelty and J. Hagmann, “The politics of security and technology in Switzerland,” *Swiss Political Science Review*, vol. 27, no. 1, pp. 128–138, 2021. doi: 10.1111/spsr.12430.

[82] R. Binns, “On the apparent conflict between individual and group fairness,” in *Proc. 2020 Conference on Fairness, Accountability, and Transparency*, 2020, pp. 514–524. doi: 10.1145/3351095.3372864.

[83] S. Dadson *et al.*, “A restatement of the natural science evidence concerning catchment-based ‘natural’ flood management in the UK,” *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 473, no. 2199, 2017. doi: 10.1098/rspa.2016.0706.

[84] A. DaViera, M. Uriostegui, A. Gottlieb, and O. Onyeka, “Risk, race, and predictive policing: A critical race theory analysis of the strategic subject list,” *American Journal of Community Psychology*, vol. 73, no. 1–2, pp. 91–103, 2023. doi: 10.1002/ajcp.12671.

[85] P. Ugwudike, “AI audits for assessing design logics and building ethical systems: The case of predictive policing algorithms,” *AI and Ethics*, vol. 2, no. 1, pp. 199–208, 2021. doi: 10.1007/s43681-021-00117-5.

[86] T. Hai, A. M. Yusof, S. Ismail, and L. Wei, “A conceptual study of key barriers in construction project coordination,” *Journal of Organizational Management Studies*, pp. 1–14, 2012. doi: 10.5171/2012.795679.

[87] A. J. Beck and A. Blumstein, “Racial disproportionality in U.S. state prisons: Accounting for the effects of racial and ethnic differences in criminal involvement, arrests, sentencing, and time served,” *Journal of Quantitative Criminology*, vol. 34, no. 3, pp. 853–883, 2017. doi: 10.1007/s10940-017-9357-6.

[88] A. Adekugbe and C. Ibeh, “Optimizing dental health equity: Integrating business analytics and program management for underserved populations in the U.S.,” *International Journal of Frontiers in Life Science Research*, vol. 6, no. 2, pp. 011–020, 2024. doi: 10.53294/ijflsr.2024.6.2.0031.

[89] A. Chapman, P. Grylls, P. Ugwudike, D. Gammack, and J. Ayling, “A data-driven analysis of the interplay between criminological theory and predictive policing algorithms,” in *Proc. 2022 ACM Conference*, pp. 36–45, 2022. doi: 10.1145/3531146.3533071.

[90] B. Bradford, K. Hohl, J. Jackson, and S. MacQueen, “Obeying the rules of the road: Procedural justice, social identity and normative compliance,” *SSRN Electronic Journal*, 2014. doi: 10.2139/ssrn.2530633.

[91] W. Perry, B. McInnis, C. C. Price, S. Smith, and J. S. Hollywood, *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*. Santa Monica, CA, USA: RAND Corporation, 2013. doi: 10.7249/RR233.

[92] B. Gogov, “Predictive policing,” 2023, pp. 207–215. doi: 10.20544/icp.8.1.23.p20.

[93] S. Hossain, A. Abtahee, I. Kashem, M. Hoque, and I. Sarker, “Crime prediction using spatio-temporal data,” 2020, pp. 277–289. doi: 10.1007/978-981-15-6648-6_22.

[94] S. Sejal, “Crime predictive model for hotspot mapping,” *International Journal for Research in Applied Science and Engineering Technology*, vol. 12, no. 6, pp. 1306–1311, 2024. doi: 10.22214/IJRASET.2024.63305.

[95] M. Feng *et al.*, “Big data analytics and mining for effective visualization and trends forecasting of crime data,” *IEEE Access*, vol. 7, pp. 106111–106123, 2019. doi: 10.1109/ACCESS.2019.2930410.

[97] A. I. Okunlola, T. Ewulo, K. E. Owolabi, and J. O. Okunlola, “Perceived influence of trees and greens in open spaces on microclimate: A case study of the Federal University of Technology, Akure,” *Progressive Horticulture*, vol. 49, no. 1, p. 86, 2017. doi: 10.5958/2249-5258.2017.00019.7.

[98] D. Weisburd, K. Petersen, C. Telep, and S. Fay, “Can increasing preventive patrol in large geographic areas reduce crime? A systematic review and meta-analysis,” *Criminology & Public Policy*, vol. 23, no. 3, pp. 721–743, 2024. doi: 10.1111/1745-9133.12665.

[99] A. A. Braga, D. M. Kennedy, E. Waring, and A. M. Piehl, “Problem-oriented policing, deterrence, and youth violence: An evaluation of Boston's Operation Ceasefire,” in *Gangs*, 2017, pp. 513–543. doi: 10.4324/9781351157803-28.

[100] S. Afzal, R. Maciejewski, and D. S. Ebert, “Visual analytics decision support environment for epidemic modeling and response evaluation,” in *Proc. IEEE Conference on Visual Analytics Science and Technology (VAST)*, 2011, pp. 191–200. doi: 10.1109/VAST.2011.6102457.

[101] R. Dasgupta, S. Mekala, F. Jaigirdar, A. Anwar, and L. Chang, “Unlocking Australia's AI usage in law enforcement from human involvement perspective: A systematic literature review,” 2025. doi: 10.21203/rs.3.rs-5933420/v1.

[102] E. Groff, J. Ratcliffe, C. Haberman, E. Sorg, N. Joyce, and R. Taylor, “Does what police do at hot spots matter? The Philadelphia policing tactics experiment,” *Criminology*, vol. 53, no. 1, pp. 23–53, 2014. doi: 10.1111/1745-9125.12055.

[103] G. Bediroglu and H. Colak, “Analysis and visualization of crime data using GIS technology: Understanding crime patterns and distribution,” *Journal of Geodesy and Geoinformation*, vol. 10, no. 2, pp. 151–163, 2023. doi: 10.9733/jgg.2023r0011.e.

[104] M. Adepeju, G. Rosser, and T. Cheng, “Novel evaluation metrics for sparse spatio-temporal point process hotspot predictions — A crime case study,” *International Journal of Geographical Information Science*, vol. 30, no. 11, pp. 2133–2154, 2016. doi: 10.1080/13658816.2016.1159684.

[105] X. Ye, L. Duan, and Q. Peng, “Spatiotemporal prediction of theft risk with deep inception-residual networks,” *Smart Cities*, vol. 4, no. 1, pp. 204–216, 2021. doi: 10.3390/smartcities4010013.

[106] A. G. Ferguson, “Predictive policing theory,” in *The Cambridge Handbook of Policing in the United States*, 2019, pp. 491–510. doi: 10.1017/9781108354721.025.

[107] A. Tak, “The role of cloud computing in modernizing healthcare IT infrastructure,” 2023, pp. 1–7. doi: 10.47363/JAICC/2023(2)151.

[108] D. Susser, “Predictive policing and the ethics of preemption,” in *Predictive Policing and the Ethics of Preemption*, 2021, pp. 268–292. doi: 10.18574/nyu/9781479803729.003.0013.

[109] A. Jiao, “Police strategies against gun crimes in the United States,” *Humanities and Social Sciences*, vol. 7, no. 1, p. 28, 2019. doi: 10.11648/j.hss.20190701.14.

[110] J. O. Okpako, *Predicting the Probability of Crime Related Danger in Los Angeles*. Rochester, NY, USA: Rochester Institute of Technology, 2025.

[111] S. Cheng, *Predictive Policing in Practice: A Case Study of Chicago's Strategic Subject List*, 2025.

[112] S. Egbert and M. Leese, *Criminal Futures*. London, U.K.: Routledge, 2020. doi: 10.4324/9780429328732.

[113] V. M. Southerland, “The intersection of race and algorithmic tools in the criminal legal system,” *Maryland Law Review*, vol. 80, p. 487, 2020.

[114] A. P. Felt *et al.*, “Improving SSL warnings: Comprehension and adherence,” in *Proc. 33rd Annual ACM Conference on Human Factors in Computing Systems*, Apr. 2015, pp. 2893–2902.

[115] G. B. Mensah, “Artificial intelligence and ethics: A comprehensive review of bias mitigation, transparency, and accountability in AI systems,” Preprint, vol. 10, no. 1, pp. 1–10, 2023.

[116] L. Harris, *Reimagining Public Safety Through Predictive Technologies*, 2023.

[117] D. Weisburd, A. A. Braga, E. R. Groff, and A. Wooditch, “Can hot spots policing reduce crime in urban areas? An agent-based simulation,” *Criminology*, vol. 55, no. 1, pp. 137–173, 2017.

[118] A. Meijer and M. Wessels, “Predictive policing: Review of benefits and drawbacks,” *International Journal of Public Administration*, vol. 42, no. 12, pp. 1031–1039, 2019.

[119] K. C. Gonugunta, “Role of analytics in offender management systems,” *The Computertech*, pp. 27–36, 2018.

[120] J. S. Kumar, M. Amiruzzaman, A. A. Bhuiyan, and D. Bhati, "Predictive analytics in law enforcement: Unveiling patterns in NYPD crime through machine learning and data mining," *Research Briefs on Information and Communication Technology Evolution*, vol. 10, pp. 36–59, 2024.

[121] S. J. Cranmer and B. A. Desmarais, "What can we learn from predictive modeling?," *Political Analysis*, vol. 25, no. 2, pp. 145–166, 2017.

[122] A. G. Ferguson, "Policing predictive policing," *Washington University Law Review*, vol. 94, p. 1109, 2016.

[123] M. Yang, S. C. Wong, and J. Coid, "The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools," *Psychological Bulletin*, vol. 136, no. 5, p. 740, 2010.

[124] B. Mills, *Firearm-Related Morbidity and Mortality by Injury Intent: Analysis of Medical, Criminal, and Vital Records in Seattle, WA*, Doctoral dissertation, 2017.

[125] G. O. Mohler *et al.*, "Randomized controlled field trials of predictive policing," *Journal of the American Statistical Association*, vol. 110, no. 512, pp. 1399–1411, 2015.

[126] S. Oosterloo and G. van Schie, "The politics and biases of the 'crime anticipation system' of the Dutch police," in *Proc. International Workshop on Bias in Information, Algorithms, and Systems*, Sheffield, U.K., 2018, pp. 30–41.

[127] S. Shukla, P. K. Jain, C. R. Babu, and R. Pamula, "A multivariate regression model for identifying, analyzing and predicting crimes," *Wireless Personal Communications*, vol. 113, no. 4, pp. 2447–2461, 2020.

[128] A. K. Meera and M. D. Jayakumar, "Determinants of crime in a developing country: A regression model," *Applied Economics*, vol. 27, no. 5, pp. 455–460, 1995.

[129] W. Safat, S. Asghar, and S. A. Gillani, "Empirical analysis for crime prediction and forecasting using machine learning and deep learning techniques," *IEEE Access*, vol. 9, pp. 70080–70094, 2021.

[130] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nature Machine Intelligence*, vol. 1, no. 5, pp. 206–215, 2019.

[131] B. Rajiv, K. Roshan, and V. Sankaradass, "Assessing the impact of bias in training data on the fairness and equity of predictive policing models," in *Proc.*

2023 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI), 2023, pp. 1–5.

[132] A. Malik, R. Maciejewski, S. Towers, S. McCullough, and D. S. Ebert, "Proactive spatiotemporal resource allocation and predictive visual analytics for community policing and law enforcement," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 1863–1872, 2014.

[133] E. L. Piza, *Police Technologies for Place-Based Crime Prevention*, 2019.

[134] B. J. Jefferson, "Predictable policing: Predictive crime mapping and geographies of policing and race," *Annals of the American Association of Geographers*, vol. 108, no. 1, pp. 1–16, 2018.

[135] T. T. Khoei and A. Singh, "A survey of emotional artificial intelligence and crimes: Detection, prediction, challenges and future direction," *Journal of Computational Social Science*, vol. 7, no. 3, pp. 2359–2402, 2024.

[136] P. Sarzaeim, Q. H. Mahmoud, A. Azim, G. Bauer, and I. Bowles, "A systematic review of using machine learning and natural language processing in smart policing," *Computers*, vol. 12, no. 12, p. 255, 2023.

[137] S. Oosterloo, *Predicting Crime Through Data: Analysis of the Data Assemblage of the Dutch Crime Anticipation System*, Master's thesis, 2021.

[138] J. Pinney, V. Bentotahewa, and M. Tomlinson, "Exploring applications and implications of big data predictive analytics in policing cyberspace," in *Navigating the Intersection of Artificial Intelligence, Security, and Ethical Governance: Sentinels of Cyberspace*, Cham, Switzerland: Springer Nature, 2024, pp. 1–18.

[139] R. Nivetha and C. Meenakshi, "Predicting crime hotspots using machine learning techniques."

[140] L. Auret and C. Aldrich, "Interpretation of nonlinear relationships between process variables by use of random forests," *Minerals Engineering*, vol. 35, pp. 27–42, 2012.

[141] N. Shiode, S. Shiode, H. Nishi, and K. Hino, "Seasonal characteristics of crime: An empirical investigation of the temporal fluctuation of different types of crime in London," *Computational Urban Science*, vol. 3, no. 1, p. 19, 2023.

[142] O. Olukoya, "Time series-based quantitative risk models: Enhancing accuracy in forecasting and risk assessment," *International Journal of Computer Applications Technology and Research*, vol. 12, no. 11, pp. 29–41, 2023.

[143] N. Rane, S. P. Choudhary, and J. Rane, “Ensemble deep learning and machine learning: Applications, opportunities, challenges, and future directions,” *Studies in Medical and Health Sciences*, vol. 1, no. 2, pp. 18–41, 2024.

[144] I. Mugari and E. E. Obioha, “Predictive policing and crime control in the United States of America and Europe: Trends in a decade of research and the future of predictive policing,” *Social Sciences*, vol. 10, no. 6, p. 234, 2021.

[145] T. C. Hart, “Hot spots of crime: Methods and predictive analytics,” in *Geographies of Behavioural Health, Crime, and Disorder: The Intersection of Social Problems and Place*, Cham, Switzerland: Springer, 2020, pp. 87–103.

[146] A. Adjekum, M. Ienca, and E. Vayena, “What is trust? Ethics and risk governance in precision medicine and predictive analytics,” *OMICS: A Journal of Integrative Biology*, vol. 21, no. 12, pp. 704–710, 2017.

[147] K. Morik and M. Scholz, “The MiningMart approach to knowledge discovery in databases,” in *Intelligent Technologies for Information Analysis*, Berlin, Germany: Springer, 2004, pp. 47–65.

[148] A. M. Shimaoka, R. C. Ferreira, and A. Goldman, “The evolution of CRISP-DM for data science: Methods, processes and frameworks,” *SBC Reviews on Computer Science*, vol. 4, no. 1, pp. 28–43, 2024.

[149] S. Tang, B. He, C. Yu, Y. Li, and K. Li, “A survey on Spark ecosystem: Big data processing infrastructure, machine learning, and applications,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 1, pp. 71–91, 2020.

[150] E. Ahishakeye, *A Business Intelligence System to Support Crime Management in Law Enforcement Agencies: A Case of Uganda Police Force*, Doctoral dissertation, University of Nairobi, 2017.

[151] V. Pillai, “Techniques for processing and analyzing large data sets using big data analytics,” SSRN Electronic Journal, 2024.

[152] B. Thalheim, *Entity-Relationship Modeling: Foundations of Database Technology*. Berlin, Germany: Springer, 2013.

[153] E. F. Codd, “A relational model of data for large shared data banks,” *Communications of the ACM*, vol. 13, no. 6, pp. 377–387, 1970.

[155] D. Calvanese, G. De Giacomo, M. Lenzerini, D. Nardi, and R. Rosati, “Data integration in data warehousing,” *International Journal of Cooperative Information Systems*, vol. 10, no. 3, pp. 237–271, 2001.

[156] U. M. Butt, S. Letchmunan, F. H. Hassan, M. Ali, A. Baqir, and H. H. R. Sherazi, "Spatio-temporal crime hotspot detection and prediction: A systematic literature review," *IEEE Access*, vol. 8, pp. 166553–166574, 2020.

[157] L. Barao and C. Mastroianni, "Creating solvability with real-time crime centers (RTCCs): Impacts on homicide and shooting investigations," *Police Quarterly*, vol. 28, no. 3, pp. 340–370, 2025.

[158] J. M. Modise, *Balancing Safety and Justice: The Ethics of Predictive Policing*.

[159] F. Coudert, D. Butin, and D. Le Métayer, "Body-worn cameras for police accountability: Opportunities and risks," *Computer Law & Security Review*, vol. 31, no. 6, pp. 749–762, 2015.

[160] R. Hadia *et al.*, "Strategic research design: Navigating methodological challenges for academic excellence," *Emerging Trends in Medical Sciences*, vol. 37, 2024.

[161] P. O. H. Wikström and R. J. Sampson, "Social mechanisms of community influences on crime and pathways in criminality," in *Causes of Conduct Disorder and Juvenile Delinquency*, vol. 118, pp. 152–178, 2003.

[162] N. Mays, C. Pope, and J. Popay, "Systematically reviewing qualitative and quantitative evidence to inform management and policy-making in the health field," *Journal of Health Services Research & Policy*, vol. 10, suppl. 1, pp. 6–20, 2005.

[163] A. Sergi, *From Mafia to Organised Crime: A Comparative Analysis of Policing Models*. Cham, Switzerland: Springer, 2017.

[164] K. C. Simmons, "The politics of policing: Ensuring stakeholder collaboration in the federal reform of local law enforcement agencies," *Journal of Criminal Law and Criminology*, pp. 489–546, 2008.

[165] D. Nagpal, I. Kornerup, and M. P. Gibson, "Mixed-method research: A basic understanding," *CODS Journal of Dentistry*, vol. 12, no. 1, pp. 11–16, 2021.

[166] D. E. Apollonio and L. A. Bero, "Interpretation and use of evidence in state policymaking: A qualitative analysis," *BMJ Open*, vol. 7, no. 2, 2017. doi: 10.1136/bmjopen-2016-012738.

[167] Y. Li, T. Wu, P. Hui, D. Jin, and S. Chen, "Social-aware D2D communications: Qualitative insights and quantitative analysis," *IEEE Communications Magazine*, vol. 52, no. 6, pp. 150–158, 2014.

[168] B. Duffy, R. Wake, T. Burrows, and P. Bremner, “Closing the gaps—crime and public perceptions,” *International Review of Law, Computers & Technology*, vol. 22, no. 1–2, pp. 17–44, 2008.

[169] L. A. Curry, I. M. Nembhard, and E. H. Bradley, “Qualitative and mixed methods provide unique contributions to outcomes research,” *Circulation*, vol. 119, no. 10, pp. 1442–1452, 2009.

[170] K. Hendren, K. Newcomer, S. K. Pandey, M. Smith, and N. Sumner, “How qualitative research methods can be leveraged to strengthen mixed methods research in public policy and public administration,” *Public Administration Review*, vol. 83, no. 3, pp. 468–485, 2023.

[171] J. Goldfedder, *Building a Data Integration Team: Skills, Requirements, and Solutions for Designing Integrations*. New York, NY, USA: Apress, 2020.

[172] S. S. Nudurupati, S. Tebboune, P. Garengo, R. Daley, and J. Hardman, “Performance measurement in data intensive organisations: Resources and capabilities for decision-making process,” *Production Planning & Control*, vol. 35, no. 4, pp. 373–393, 2024.

[173] A. N. Prasad, “Data quality and preprocessing,” in *Introduction to Data Governance for Machine Learning Systems: Fundamental Principles, Critical Practices, and Future Trends*, Berkeley, CA, USA: Apress, 2024, pp. 109–223.

[174] S. S. Alhabsi, “Ethical considerations in obtaining informed consent in research participation,” *International Journal of Educational Contemporary Explorations*, vol. 1, no. 1, pp. 22–32, 2024.

[175] S. Roberts, “Rebuilding trust in a divided community: An integrated approach,” *Pepperdine Dispute Resolution Law Journal*, vol. 24, p. 461, 2024.

[176] W. M. Lim, “What is qualitative research? An overview and guidelines,” *Australasian Marketing Journal*, vol. 33, no. 2, pp. 199–229, 2025.

[177] M. N. Ismail *et al.*, “Quantitative insights and challenges in big data from a statistical perspective,” *Journal of Ecohumanism*, vol. 3, no. 5, pp. 290–307, 2024.

[178] P. Cash, O. Isaksson, A. Maier, and J. Summers, “Sampling in design research: Eight key considerations,” *Design Studies*, vol. 78, p. 101077, 2022.

[179] A. Ahmad *et al.*, “Unveiling urban violence crime in the state of Selangor, Kuala Lumpur and Putrajaya: A spatial-temporal investigation of violence crime in

Malaysia's key cities," *Cogent Social Sciences*, vol. 10, no. 1, 2024. doi: 10.1080/23311886.2024.2347411.

[180] L. S. Chiew, S. Amerudin, and Z. M. Yusof, "A spatial analysis of the relationship between socio-demographic characteristics with burglar behaviours on burglary crime," in *IOP Conference Series: Earth and Environmental Science*, vol. 540, no. 1, p. 012050, 2020.

[181] G. Song *et al.*, "Residents, employees and visitors: Effects of three types of ambient population on theft on weekdays and weekends in Beijing, China," *Journal of Quantitative Criminology*, vol. 39, no. 2, pp. 385–423, 2023.

[182] M. Rty, M. Kuronen, M. Myllymäki, A. Kangas, K. Mäkisara, and J. Heikkinen, "Comparison of the local pivotal method and systematic sampling for national forest inventories," *Forest Ecosystems*, vol. 7, no. 1, p. 54, 2020.

[183] V. Ceccato and I. Ioannidis, "Using remote sensing data in urban crime analysis: A systematic review of English-language literature from 2003 to 2023," *International Criminal Justice Review*, vol. 35, no. 2, pp. 102–122, 2025.

[184] M. I. Sameen, B. Pradhan, D. T. Bui, and A. M. Alamri, "Systematic sample subdividing strategy for training landslide susceptibility models," *Catena*, vol. 187, p. 104358, 2020.

[185] A. Godfroid and S. Andringa, "Uncovering sampling biases, advancing inclusivity, and rethinking theoretical accounts in second language acquisition: Introduction to the special issue SLA for all?," *Language Learning*, vol. 73, no. 4, pp. 981–1002, 2023.

[186] A. Ahmad, T. Masron, A. S. M. Ali, Y. Kimura, and S. N. Junaini, "Demographic dynamics and urban property crime: A linear regression analysis in Kuala Lumpur and Putrajaya (2015–2020)," *Planning Malaysia*, vol. 22, 2024.

[187] R. T. Kneusel, *The Art of Randomness: Randomized Algorithms in the Real World*. San Francisco, CA, USA: No Starch Press, 2024.

[188] E. Thomann and M. Maggetti, "Designing research with qualitative comparative analysis (QCA): Approaches, challenges, and tools," *Sociological Methods & Research*, vol. 49, no. 2, pp. 356–386, 2020.

[189] J. R. Hipp and S. A. Williams, "Advances in spatial criminology: The spatial scale of crime," *Annual Review of Criminology*, vol. 3, no. 1, pp. 75–95, 2020.

[190] M. S. Redzuan *et al.*, “Assessing Malaysia's urban security through emerging hot spot analysis (EHSA): A spatiotemporal investigation of burglary patterns within police jurisdictions of Selangor, Kuala Lumpur, and Putrajaya, Malaysia,” *International Journal of Geoinformatics*, vol. 21, no. 3, 2025.

[199] D. Patel *et al.*, “Contemporary practice and considerations for real-world data source identification and feasibility assessment,” *Pharmacoepidemiology and Drug Safety*, vol. 33, no. 9, p. e5862, 2024.

[200] N. Hamilton-Smith and S. Mackenzie, “The geometry of shadows: A critical review of organised crime risk assessments,” *Policing and Society*, vol. 20, no. 3, pp. 257–279, 2010.

[201] A. J. Zoutendijk, “Organised crime threat assessments: A critical review,” *Crime, Law and Social Change*, vol. 54, no. 1, pp. 63–86, 2010.

[202] A. Adekugbe and C. Ibeh, “Navigating ethical challenges in data management for U.S. program development: Best practices and recommendations,” *International Journal of Management & Entrepreneurship Research*, vol. 6, no. 4, pp. 1023–1033, 2024.

[203] O. Alabi and T. Bukola, “Introduction to descriptive statistics,” *IntechOpen*, 2023.

[204] B. Nadeak and L. Naibaho, “The effectiveness of mastery learning technique on improving students’ ability in completing English national examination,” *International Journal of Research – Granthaalayah*, vol. 8, no. 2, pp. 57–62, 2020.

[205] K. Najarian and R. Abdi, “The relationship between abnormal personality dimensions and elderly interpersonal problems,” *Salmand*, vol. 13, no. 10, 2017.

[206] W. Zhu, H. Zheng, H. Liao, W. Li, and J. Luo, “Learning bias-invariant representation by cross-sample mutual information minimization,” *arXiv preprint*, 2021.

[207] A. Jain *et al.*, “Overview and importance of data quality for machine learning tasks,” in *Proc. 26th ACM SIGKDD Int. Conf. Knowledge Discovery & Data Mining*, 2020, pp. 3561–3562.

[208] I. N. Pratama, M. Dachyar, and N. R. Pratama, “Optimization of resource allocation and task allocation with project management information systems in information technology companies,” *TEM Journal*, vol. 12, no. 3, p. 1814, 2023.

[209] S. Sivasamy, “Sample size considerations in research,” *Endodontology*, vol. 35, no. 4, pp. 304–308, 2023.

[210] M. E. Wende, M. R. U. Meyer, C. G. Abildso, K. Davis, and A. T. Kaczynski, “Urban–rural disparities in childhood obesogenic environments in the United States: Application of differing rural definitions,” *The Journal of Rural Health*, vol. 39, no. 1, pp. 121–135, 2023.

[211] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to Linear Regression Analysis*, 6th ed. Hoboken, NJ, USA: John Wiley & Sons, 2021.

[212] A. E. Streetman, A. K. Lewis, E. L. Rogers, K. M. Heinrich, and J. A. DeBlauw, “Anticipatory anxiety, familiarization, and performance: Finding the sweet spot to optimize high-quality data collection and minimize subject burden,” *European Journal of Investigation in Health, Psychology and Education*, vol. 12, no. 9, pp. 1349–1357, 2022.

[213] M. Khosravi, Z. Zare, S. M. Mojtabaeian, and R. Izadi, “Artificial intelligence and decision-making in healthcare: A thematic analysis of a systematic review of reviews,” *Health Services Research and Managerial Epidemiology*, vol. 11, 2024.

[214] M. Kahng, S. Navathe, J. Stasko, and D. Chau, “Interactive browsing and navigation in relational databases,” *Proceedings of the VLDB Endowment*, vol. 9, no. 12, pp. 1017–1028, 2016.

[215] S. Al-Fedaghi, “Conceptual data modeling: Entity-relationship models as thinging machines,” *arXiv preprint*, 2021.

[216] P. Das, A. K. Das, J. Nayak, and D. Pelusi, “A framework for crime data analysis using relationship among named entities,” *Neural Computing and Applications*, vol. 32, no. 12, pp. 7671–7689, 2020.

[217] E. Männistö, *Building a Simple PHP Framework*, 2023.

[218] K. Král and L. Szomorová, “The possibilities of using chosen jQuery JavaScript components in creating interactive maps,” *Geomatics Landmanagement and Landscape*, no. 2, pp. 45–54, 2015.

[219] C. Carthen *et al.*, “A novel spatial data pipeline for orchestrating Apache NiFi/MiNiFi,” *International Journal of Software Innovation*, vol. 12, no. 1, pp. 1–14, 2023.

[220] J. Zhang, “Design and implementation of pedestrian detection system based on video image,” 2023.

[221] J. K. A. Oluwaferanmi, “Automating ETL pipelines using artificial intelligence: Transforming legacy data integration systems into intelligent data workflows,” 2025.

[222] A. Thirunagalingam, “Combining AI paradigms for effective data imputation: A hybrid approach,” *International Journal of Transformations in Business Management*, vol. 14, no. 1, pp. 10–37648, 2024.

[223] K. J. Olowe, N. L. Edoh, S. J. C. Zouo, and J. Olamijuwon, “Conceptual review on the importance of data visualization tools for effective research communication,” *International Journal of Engineering Research and Development*, vol. 20, no. 11, pp. 1259–1268, 2024.

[224] M. Kansara, “Advancements in cloud database migration: Current innovations and future prospects for scalable and secure transitions,” 2024.

[225] Y. Wang, H. Wang, and Y. Cao, “Comprehensive review of storage optimization techniques in blockchain systems,” *Applied Sciences*, vol. 15, no. 1, p. 243, 2024.

[226] O. Ilori, N. T. Nwosu, and H. N. N. Naiho, “Advanced data analytics in internal audits: A conceptual framework for comprehensive risk assessment and fraud detection,” *Finance & Accounting Research Journal*, vol. 6, no. 6, pp. 931–952, 2024.

[227] M. Ahmad and S. Wilkins, “Purposive sampling in qualitative research: A framework for the entire journey,” *Quality & Quantity*, vol. 59, no. 2, pp. 1461–1479, 2025.

[228] M. R. Maidin, M. F. Hussin, N. A. Z. Zakaria, S. A. E. A. Rahim, and K. Samsudin, “The need for integrated crime risk assessment systems using data-driven strategies for crime prevention in Malaysia,” *International Journal of Academic Research in Business and Social Sciences*, vol. 14, no. 11, pp. 996–1005, 2024.

[229] S. Murni, M. Ahkam, D. Sari, I. Ichsan, and E. Elfisa, “Sewing service order information web-based (a case study at Iendandy Modiste),” *Jurnal Inotera*, vol. 8, no. 2, pp. 272–283, 2023.

[230] R. F. Marasambessy, “Information technology-based law enforcement in increasing public trust in the police,” *Gema Wiralodra*, vol. 14, no. 2, pp. 790–798, 2023.

APPENDICES

APPENDIX 1

Letter of Application for Conducting Research and Development of the ICRAS System for the Royal Malaysian Police

www.uitm.edu.my

UNIVERSITI TEKNOLOGI MARA | College of Engineering

MOHD RAZIF BIN MAIDIN
FAKULTI KEJURUTERAAN ELEKTRIK
UITM SHAH ALAM
40100, SHAH ALAM
SELANGOR.

Rujukan: 2021405254

Tarikh: APR 2022

KETUA POLIS DAERAH SHAH ALAM,
IBU PEJABAT POLIS DAERAH SHAH ALAM,
PERSIARAN KAYANGAN,
SEKSYEN 9,
40100, SHAH ALAM
SELANGOR.

YDH Tuan,

KERJASAMA MELAKSANAKAN PENYELIDIKAN PENGURUSAN JENAYAH DI MALAYSIA ANTARA PDRM - UiTM

Dengan segala hormatnya izinkan saya menarik perhatian YDH Tuan berkenaan perkara di atas.

2. Dimaklumkan bahawa Kolej Kejuruteraan Universiti Teknologi Mara (UiTM) Shah Alam sedang melaksanakan penyelidikan berkaitan pengurusan jenayah di Malaysia oleh Polis Diraja Malaysia. Objektif

3. Untuk makluman YDH Tuan juga, penyelidikan ini dijalankan oleh pelajar Penyelidikan Dokor Falsafah dibawah Kolej Kejuruteraan (KK) UiTM Shah Alam seperti butiran:

3.1.	Pangkat / No Badan / Nama Pegawai	:	Insp G/21603 Mohd Razif Bin Maidin
3.2.	Tempat Tugas	:	Cawangan Khas Bahagian E1 IPK Shah Alam
3.3.	Nama Pensyarah Penyelia	:	Prof. Madya Ts. Dr. Mohamad Fahmi Bin Hussin @ Mohamad
3.4.	No telefon	:	

4. Bagi tujuan tersebut, pihak kami memohon mengadakan satu mesyuarat koordinasi diantara pihak KK UiTM Shah Alam dan YDH Ketua Polis Daerah Shah Alam bagi mengadakan perbincangan berkaitan.

5. Dikemukakan untuk pertimbangan dan kelulusan dari YDH Tuan.

Sekian, terima kasih.

(MOHD RAZIF BIN MAIDIN) INSP
INSP E1A (2) SEL

MRM / MRM..

APPENDIX 2

Official letter from the Royal Malaysian Police (RMP) stating that the development of the ICRAS system is promising and should be pursued as an initiative within the RMP



POLIS DIRAJA MALAYSIA
IBU PEJABAT POLIS KONTINJEN SELANGOR
40912 SHAH ALAM, SELANGOR

Telefon: 03- 55145004

Faks: 03-603-5510537

Rujukan: 12/2021

Tarikh : 25 MAC 2022

KOLEJ KEJURUTERAAN,
LEVEL 9, MENARA 2,
KOMPLEKS KEJURUTERAAN,
TUANKU ABDUL HALIM MU'ADZAM SHAH,
UNIVERSITI TEKNOLOGI MARA,
40450, SHAH ALAM,
SELANGOR.

Prof / Prof Madya / Ir / Ts / Dr ,

PEMBANGUNAN PERISIAN APPLIKASI ANALISIS RISIKO JENAYAH BAGI KEGUNAAN JABATAN SIASATAN JENAYAH (JSJ)

Dengan segala hormatnya perkara di atas adalah dirujuk.

2. Dimaklumkan bahawa salah seorang Pegawai Kanan Polis dari Ibu Pejabat Polis Kontinjen SELANGOR sedang melanjutkan pengajian dalam peringkat Doktor Falsafah di Kolej Kejuruteraan , UITM Shah Alam. Butiran adalah seperti berikut:-

2.1.	Pangkat / No Badan / Nama Pegawai	: Insp G/21603 Mohd Razif Bin Maidin
2.2.	Tempat Tugas	: Cawangan Khas IPK Shah Alam
2.3.	Nama Pensyarah Penyelia	: Prof. Madya Ts. Dr. Mohamad Fahmi Bin Hussin @ Mohamad

...2/-

3. Skop pengajian pegawai ini adalah meliputi analisis risiko jenayah , keselamatan dan kesihatan Pegawai Penyiasat yang akan menghasilkan satu applikasi dalam peranti mudah alih bagi membantu Pegawai Penyiasat dalam melaksanakan tugasan di lapangan dengan lebih komprehensif.

4. Jabatan berpendapat bahawa ini adalah merupakan satu perkara yang baru dan berpotensi untuk dimajukan. Perkara ini juga sejajar dengan hasrat kerajaan bagi mendepani Revolusi Industri 4.0. Platform ini juga dilihat sebagai satu inisiatif pemampanan Jabatan Siasatan Jenayah.

Sekian, terima kasih.

WAN FAUZI BIN WAN NAWANG) SUPT

PEGAWAI TURUS CAWANGAN KHAS SELANGOR
b.p. KETUA CAWANGAN KHAS SELANGOR
b.p. KETUA POLIS SELANGOR

WFWN / MRM..

APPENDIX 3

Minutes of the meeting during the 3rd ICRAS Coordination Meeting

IBU PEJABAT POLIS KONTINJEN SELANGOR,
POLIS DIRAJA MALAYSIA,
PETI SURAT 7036,
40912 SHAH ALAM,
SELANGOR.

TEL : 03-55145222
FAX : 03-55146558

'Polis Dan Masyarakat Berpisah Tiada'

Ruj : (PR) 33/1/3
Tarikh : 18 MAC 2024

MINIT MESYUARAT SISTEM ICRAS
BIL : 03 / 2024

Tarikh : **11 MAC 2024 (Isnin)**
Masa : **9.40 – 12.20 tgh hari**
Tempat : **Bilik Mesyuarat GTR, Tingkat 1, Menara 1, IPK Selangor**

KEHADIRAN

1.	SAC Syamsul Bin Hailin	:	Pegawai Turus E1 SELANGOR
2.	ACP Mohd Khuzaimie bin Mahamud	:	T/PTE1 Selangor
3.	ACP Zulkifli Bin Adamsah	:	Wakil KJPJKK Selangor
4.	ACP Ahmad Jefferi Bin Abdullah	:	KJSJN Selangor
5.	ACP Mohamed Izzani Bin Abdul Halim	:	Wakil Pemangku KCK Selangor
6.	ACP Muhammad Yazid Bin Muhammad Yew	:	KJSJK Selangor
7.	SUPT Azman Bin Shari'at	:	KJSPT Selangor
8.	SUPT Rudy Bin Hj Abdullah	:	KJKDN/KA Selangor
9.	DSP Mohd Kahar Bin Mantamin	:	Wakil KJLog Selangor
10.	ASP Norazmi Bin Shamsuri	:	ASP Media Selangor
11.	INSP Mohd Razif Bin Maidin	:	Pembentang
TIDAK HADIR			
1.	SAC Kumaran A/L Karuppannan	:	KJPJKK Selangor (Cuti)
2.	ACP Mazlina Binti Ab Malik	:	Pemangku KCK Selangor (Tugas)
3.	ACP Zakaria Bin Ismail	:	KJLog Selangor (Tugas)
4.	ACP Mohd Ismail Bin Muslim	:	KJIPS Selangor (Tugas)

BIL	PERKARA	TINDAKAN
1.	**PENDAHULUAN** YDH SAC Syamsul Bin Hailin, Pegawai Turus E1 Selangor selaku	**Makluman**

2.	**UCAPAN KATA ALUAN DAN SESI PEMBENTANGAN** 2.1 **<u>Pembentangan oleh INSP G/21603 berkaitan dengan data dan Sistem Appllikasi ICRAS</u>** 2.1.1. Pembentangan dan analisis data berkaitan Sistem ICRAS yang telah dikemaskini dan dipersetujui oleh SME (Subject Matter Expert). Data mewakili statistik jenayah yang berlaku di Daerah Shah Alam dalam kawasan pentadbiran IPK SELANGOR. 2.1.2 Pembentangan menyentuh tentang aturan dan carta alir sistem tetapi ianya boleh dikemaskini supaya kelihatan menarik dan lebih berinformatif. 2.1.3. Pembentangan juga menyentuh mengenai gambaran keseluruhan mengenai Sistem Laporan dan Analisis Jenayah Pintar (ICRAS), menerangkan tujuan, sejarah pembangunan, dan manfaat yang diharapkan untuk pasukan polis. Pembentangan yang komprehensif menyentuh ciri-ciri utama ICRAS seperti berikut: 2.1.4. Analisis dan Visualisasi Data: Menggunakan algoritma canggih untuk menganalisis data jenayah dan memberikan cadangan yang perlu dilakukan. 2.1.5. Polis Pencegahan: Membantu meramalkan kawasan jenayah yang berpotensi dan mengagihkan sumber dengan cekap. 2.1.6. Integrasi dengan Sistem Sedia Ada: Mengintegrasikan dengan lancar dengan pangkalan data dan rangkaian komunikasi polis semasa. 2.1.7. Tahap Akses Pengguna:Menjamin keselamatan dan privasi data dengan kebenaran akses bertingkat untuk pangkat yang berbeza. **PERBINCANGAN DAN MAKLUM BALAS** Sesi perbincangan dibuka. Poin utama yang dibangkitkan termasuk: 3.1.1. Menekankan kepentingan ketepatan data dan kemas kini	**INSP MOHD RAZIF BIN MAIDIN**

3.	3.1.2. Mengenai kebolehskalaan sistem dan persediaan masa depan. 3.1.3. Bidang yang boleh diperbaiki lagi, seperti menambah lebih banyak bahasa untuk inklusiviti.	
4.	**RINGKASAN MAKLUM BALAS** 4.1. Sambutan keseluruhan positif terhadap keupayaan ICRAS dan impaknya yang berpotensi terhadap pengurusan jenayah. 4.2. Cadangan untuk ciri tambahan dan penambahbaikan dicatat untuk pertimbangan lanjut. 4.3 **Analisis Kumpulan SME** 4.3.1. Satu kumpulan SME yang diketuai oleh Pegawai Kanan Polis berpangkat Assistant Komissioner Polis telah bersetuju dan mengesahkan keabsahan data statistik jenayah yang dikumpul dan dibentangkan. 4.3.2 Kumpulan SME juga bersetuju terhadap sistem yang dibangunkan tetapi ada beberapa yang perlu dikemaskini.	
5.	**KESIMPULAN DAN LANGKAH SETERUSNYA** 5.1.1. INSP MOHD RAZIF diminta untuk memasukkan maklum balas dan cadangan ke dalam pelan pembangunan sistem. 5.1.2 Mesyuarat susulan akan dijadualkan untuk mengkaji kemajuan dan membincangkan strategi pelaksanaan.	
6.	**LAIN- LAIN HAL** Tiada perkara berbangkit.	**Makluman**
7.	**PENANGGUHAN MESYUARAT** Pengerusi mengucapkan terima kasih kepada semua yang hadir dan menangguhkan mesyuarat pada jam 1220 hrs.	**Makluman**

Disediakan oleh :-



(Mohd Razif Bin Maidin) INSP
Pegawai Pembentang
Pembangunan Sistem Aplikasi ICRAS

Disahkan oleh :-



(Syamsul Bin Hailin) SAC
Pegawai Turus E1 SELANGOR

KP Selangor	- Salinan No.1
Timb. KP Selangor	- Salinan No.2
KJSJ Selangor	- Salinan No.3
KCK Selangor	- Salinan No.4
KJP Selangor	- Salinan No.5
KJPJKK Selangor	- Salinan No.6
KJSJK Selangor	- Salinan No.7
KJSJN Selangor	- Salinan No.8
KJStaRT Selangor	- Salinan No.9
KJSPT Selangor	- Salinan No.10
KJKDN/KA Selangor	- Salinan No.11
KJIPS Selangor	- Salinan No.12
KPD Shah Alam	- Salinan No.13
PPM Wilayah 1 Pelabuhan Klang	- Salinan No.14
Pusat Media Selangor	- Salinan No.15
ASP Kem Komandan	- Salinan No.16
PT Pentadbiran	- Salinan No.17
Pencatit minit	- Salinan No.18

APPENDIX 4

Questionnaire

Questionnaire for Integrated Crime Risk Assessment System (ICRAS) Development

This questionnaire is part of an academic research project titled **"Integrated Crime Risk Assessment System (ICRAS) for the Royal Malaysian Police (RMP)"**, which aims to design and develop a mobile and web-based application to support Investigation Officers (IOs) in assessing, managing, and monitoring crime risks more effectively.

Section 1: Demographic Information

1. What is your role in the organization?

   Mark only one oval.

   - ◯ Investigation Officer (IO)
   - ◯ Police Officer
   - ◯ Senior Police Officer
   - ◯ Other: _____

2. How long have you been working in law enforcement?

   Mark only one oval.

   - ◯ Less than 1 year
   - ◯ 1–3 years
   - ◯ 4–5 years
   - ◯ More than 5 years

3. Which department are you currently working in?

   _____

4. What is your age group?

Mark only one oval.

- ◯ Under 25
- ◯ 26-35
- ◯ 36-45
- ◯ 46-55
- ◯ 56+

Section 2: Technology and Device Usage

5. Do you use a smartphone in your daily duties?

Check all that apply.

- ☐ Yes
- ☐ No

6. What type of smartphone do you use?

Mark only one oval.

- ◯ Android
- ◯ I-Phone
- ◯ Other: _____

7. How comfortable are you with using mobile apps for work-related tasks?

Mark only one oval.

1	2	3	4	5
◯	◯	◯	◯	◯

8. How comfortable are you with using mobile apps for work-related tasks?

Mark only one oval.

1	2	3	4	5
◯	◯	◯	◯	◯

9. What do you primarily use your mobile phone for at work?

Check all that apply.

- [ ] Communication (call/message)
- [ ] Accessing Report
- [ ] GPS/Navigation
- [ ] Data Entry

10. How often do you use mobile applications to support your law enforcement duties?

Mark only one oval.

- ◯ Daily
- ◯ Weekly
- ◯ Monthly
- ◯ Never

11. Have you ever used any crime-related mobile apps for your duties?

Check all that apply.

- [ ] Yes
- [ ] No

12. If yes, what features did you find most useful in those apps?

13. How often do you conduct crime risk assessments in your role?

Mark only one oval.

- ◯ Daily
- ◯ Weekly
- ◯ Monthly
- ◯ Occassionally

14. What tools or systems do you currently use to assess crime risk?

Check all that apply.

- ☐ Manual methods (e.g., reports, papers)
- ☐ Automated systems
- ☐ Other: _____

15. How effective do you think the current crime risk assessment methods are?

Mark only one oval.

1	2	3	4	5
◯	◯	◯	◯	◯

16. Do you think that real-time data is crucial for accurate crime risk assessment?

Check all that apply.

- ☐ Yes
- ☐ No

17. What is your biggest challenge when conducting crime risk assessments?

Section 4: ICRAS System Awareness and Feedback

18. How familiar are you with the ICRAS system?

Mark only one oval.

- ◯ Not familiar
- ◯ Somewhat familiar
- ◯ Familiar
- ◯ Very familiar

19. How useful do you think the ICRAS system could be in assessing crime risks for IOs?

Mark only one oval.

1	2	3	4	5
◯	◯	◯	◯	◯

20. Do you think an app like ICRAS would help improve decision-making in your work?

Check all that apply.

- ☐ Yes
- ☐ No

21. What features would you find most useful in a crime risk assessment app like ICRAS?

Check all that apply.

- [ ] Real-time data updates
- [ ] Predictive analytics
- [ ] Crime risk heat maps
- [ ] Reporting and visualization tools
- [ ] Crime trend forecasting
- [ ] Other: _____

22. Would you be willing to use an app like ICRAS for your daily duties?

Check all that apply.

- [ ] Yes
- [ ] No

23. What concerns or challenges do you foresee with adopting an app like ICRAS?

_____

_____

_____

_____

_____

24. Do you believe a mobile application for crime risk assessment like ICRAS will improve public safety and efficiency in your work?

Check all that apply.

- [ ] Yes
- [ ] No

25. In your opinion, is there a need for an app like ICRAS to help Investigation Officers (IOs) in their work?

Check all that apply.

- [ ] Yes
- [ ] No

26. If "Yes", please explain how it would improve your work:

27. If "Yes", please explain how it would improve your work:

Section 5: Implementation and Adoption

28. How likely are you to adopt a mobile app like ICRAS if it were integrated into your daily operations?

Mark only one oval.

- ◯ Very unlikely
- ◯ Unlikely
- ◯ Neutral
- ◯ Likely
- ◯ Very likely

APPENDIX 5

Coordination meeting, SME and training – Pilot Test



ICRAS second coordination meeting with Investigation Officers (IOs)



Coordination and feedback session with Subject Matter Expert (SME) to improve system reliability and align with IOs' needs



Training and system familiarization session with end users to support effective system adoption.



Receipt of official letter from the Royal Malaysian Police (RMP) acknowledging the potential and importance of the ICRAS development initiative.

AUTHOR'S PROFILE



Mohd Razif Bin Maidin obtained his Higher National Diploma in Electronic Engineering from Kolej Kemahiran Tinggi MARA in 2004, followed by a Bachelor's degree in Electronic Engineering Technology from the British Malaysia Institute. He later completed his Master of Science in Telecommunication and Networking Engineering at Universiti Teknologi MARA (UiTM) Shah Alam in 2010. Currently, he is pursuing a PhD in Electrical Engineering at UiTM Shah Alam (expected completion in 2025). His doctoral work explores the development of an Integrated Crime Risk Assessment System (ICRAS) designed for the Royal Malaysian Police (RMP). The research introduces a new approach through the creation of a risk appetite index and crime risk matrix, enabling Investigation Officers (IOs) to assess and interpret crime risks more systematically. By combining ETL-based data processing with predictive analytics, his study aims to deliver a data driven, real time platform that supports informed decision making, enhances operational efficiency, and strengthens public safety management in Malaysia.

LIST OF PUBLICATION:

Maidin, M. R., Hussin, M. F., Zakaria, N. A. Z., AB, R. S., & SAMSUDIN, K. (2024). A Comprehensive Framework for Crime Risk Assessment in Malaysia: Integrating Data-Driven Insights for Enhanced Public

Safety. *INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN BUSINESS AND SOCIAL SCIENCES*, *14*(11).

Maidin, M. R. B., Hussin, M. F., Zakaria, N. A. Z., Ab. Rahim, S. A., & Samsudin, K. (2024). A comprehensive framework for crime risk assessment in Malaysia: Integrating data-driven insights for enhanced public safety. *International Journal of Academic Research in Business and Social Sciences, 14*(11), 996-1005. https://hrmars.com/index.php/papers/detail/MAJESS/23600/The-Need-for-Integrated-Crime-Risk-Assessment-Systems-Using-Data-Driven-Strategies-for-Crime-Prevention-in-Malaysia

Maidin, M. R., Hussin, M. F., Rashid, N. E. A., & Yiliatiningtyas, S. (2024). The impact of payload on radar detection: A comprehensive RCS analysis of drones operating in 2.4 GHz and 5.2 GHz bands. *Applied Sciences and Engineering Technology, 60*(2). https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/article/view/4898

Maidin, M. R. B. Hussin, M. F, Khairilmizal S (2025). Enhancing The Validity and Reliability of The Risk Appetite Index Questionnaire for Investigation Officers in The Royal Malaysian Police. *ASIA Journal*. Submitted and presented at the 8th ASEAN Smart Grid Congress. (Under review). Expected publication in March 2026.

Masjefri, A. Z., Hussin, M. F., Maidin, M. R. B., Sulaiman, A. A., Samsudin, K., Yona, H., & Ab. Razak, N. I. (2026). Comparative analysis of radio frequency (RF) measurements between ICNIRP and MCMC standards at Grandeur Labs, Eco Grandeur. *Journal of Applied Research and Technology* (under review). Expected publication in April 2026.

Maidin, M. R., Hussin, M. F., Samsudin, K., Zakaria, N. A. Z., & Enche Abdul Rahim, S. A. (2025, December). ***Data-driven strategies for crime prevention: A focus on integrated crime risk assessment.*** Paper presented at the International Conference on Science and Social Research (CSSR) 2025, Malaysia. (Manuscript under consideration for publication in a Scopus/WoS indexed journal).

Maidin, M. R., Hussin, M. F., Samsudin, K., Zakaria, N. A. Z., & Enche Abdul Rahim, S. A. (2025). ***Strengthening Risk Appetite Assessment Tools for Operational Decision Making: A Reliability and Readability Enhancement Study within a***

Large Public Sector Organization. PaperAsia Journal. (Manuscript submitted on NOV 2025- Expected publication in MEI 2026.

Maidin, M. R., Hussin, M. F., Samsudin, K., Zakaria, N. A. Z., & Enche Abdul Rahim, S. A. (2025). ***Mathematical Modelling and Optimization of Crime Risk Prediction Using a Random Forest Based Control Framework Within The Integrated Crime Risk Assessment System (ICRAS). Barekeng Journal of Mathematics and Its Applications.*** (Manuscript submitted on NOV 2025- Expected publication on April 2025.

Publication in E - Proceeding and Presentation in Competition

Maidin, M. R. Hussin, M. F Samsudin, K (2025). *Design and development of mobile application for assessing crime risk for investigation officers (IO)*. In The International Competition on Sustainable Education (SUSED 2025), Innovation in Educational Research Section (p. 552). UITM Negeri Sembilan. E-Proceedings. [ERS2025008].

Publication by UPM Press in edited volume book

Masjefry, A. Z., Hussin, M. F., Maidin, M. R., Samsudin, K., Sulaiman, A. A., Yona, H., & AB. Razak, N. I. (2025). Comparative Analysis of Radio Frequency (RF) Measurements Between ICNIRP And MCMC Standards at Grandeur Labs, Eco Grandeur. *Upcoming Edited Volume* (pp.). UPM Press. Expected publication in Dec 2025.

Competition and Award

Silver Award – *The 4th International Competition on Sustainable Education (SUSED) 2025*

Organized by: Universiti Teknologi MARA (UiTM) Kuala Pilah

Category: Innovation in Educational Research (Student)

Innovation Title: *Design and Development of Mobile Application for Accessing Crime Risk for Investigation Officers (IO) Using ETL-Based Data Processing Framework and Predictive Analysis*

Silver Award – *The 11th International Conference on Global Education (ICGE) 2025*
Organized by: Universiti Kebangsaan Malaysia (UKM)
Innovation Title: *Design and Development of Mobile Application for Accessing Crime Risk for Investigation Officers (IO) Using ETL-Based Data Processing Framework and Predictive Analysis*