

Comparative Analysis of Data Protection Laws and Digital Privacy in Criminal Investigations

Dr. Pooja Khurana

Assistant Professor

Department of Law

Maharaja Agrasen Himalayan Garhwal University

ORCID id - <https://orcid.org/0009-0000-7767-3527>



<http://www.ejset.org/> || Vol. 2 No. 3 (2026): September Issue

Date of Submission: 05-08-2026

Date of Acceptance: 20-08-2026

Date of Publication: 04-09-2026

Abstract— Digital criminal investigations increasingly depend on smartphones, cloud platforms, biometric repositories, location records, social networks, and algorithmic tools capable of processing large volumes of personal information. This expansion of investigative capability has created a difficult legal problem: privacy legislation must protect individuals from disproportionate surveillance without preventing authorities from obtaining legitimate digital evidence. The present study develops a comparative framework for examining this tension across the European Union, United Kingdom, United States, and India, with particular attention to developments between 2021 and 2026. A central research gap is identified in the limited integration of conventional comparative data-protection analysis with the operational consequences of artificial intelligence, automated evidence triage, biometric analytics, cross-border cloud access, and machine-supported investigative decision-making. The study therefore proposes the concept of an Investigative Privacy-Proportionality Framework (IPPF) for evaluating how different legal systems regulate necessity, purpose limitation, sensitive-data processing, automated decisions, transparency, accountability, and international data access.

Keywords— Digital Privacy, Data Protection Law, Criminal Investigation, Artificial Intelligence, Digital Evidence, Law-Enforcement Data Processing

Introduction

The transformation of criminal investigation from a predominantly physical process into a data-intensive activity has altered the meaning of investigative privacy. Contemporary investigations can involve extraction of mobile-device contents, analysis of cloud accounts, geolocation histories, communications metadata, biometric identification, CCTV footage, financial transactions, online behavioural traces, Internet-of-Things records, and social-media connections. Increasingly, these materials are not examined exclusively by individual investigators. Machine-learning systems, facial-recognition tools, automated forensic platforms, entity-resolution systems, natural-language processing models, and risk-scoring mechanisms can identify correlations among thousands or millions of digital records.

The resulting investigative environment raises a fundamental legal question. Access to relevant digital information may be necessary for preventing serious crime and identifying offenders, yet large-scale acquisition and computational analysis may expose information concerning suspects,

victims, witnesses, family members, professional contacts, patients, customers, and other individuals who have no direct involvement in the offence under investigation. Consequently, the traditional legal distinction between a targeted search and general surveillance becomes increasingly difficult to maintain when investigators acquire an entire digital device or cloud account and subsequently use software to identify material considered relevant.

Data protection law therefore has an increasingly important relationship with criminal procedure. Conventional search-and-seizure rules generally determine whether investigators may obtain information, whereas modern data-protection frameworks additionally address how acquired personal information should be processed, retained, reused, transferred, secured, or subjected to automated analysis. These two regulatory layers do not always develop at the same pace. Investigative authority may expand through technological capability before detailed rules emerge concerning algorithmic analysis, secondary processing, model-generated inferences, or the retention of information belonging to individuals who are not ultimately charged.

The European legal framework provides one of the clearest examples of specialised law-enforcement data governance. Processing by competent authorities for criminal-law purposes is principally governed by Directive (EU) 2016/680, commonly known as the Law Enforcement Directive (LED), rather than the ordinary GDPR framework. The Directive contains rules addressing lawfulness, purpose limitation, data quality, sensitive personal information and automated decision-making. In particular, Article 11 restricts decisions based solely on automated processing that produce adverse legal or similarly significant effects unless authorised by law and accompanied by appropriate safeguards, including human intervention.

The United Kingdom retained a similarly differentiated structure through Part 3 of the Data Protection Act 2018. Law-enforcement processing must be lawful and fair, and processing conducted by a competent authority generally

requires a statutory or other legal basis and necessity for a law-enforcement task. Sensitive processing attracts additional safeguards, including conditions involving strict necessity in relevant circumstances. UK legislation also incorporates principles relating to specified purposes and accountability.



Fig. 1: Right to Digital Privacy

Source: <https://www.sconline.com/blog/post/2021/10/01/right-to-digital-privacy/>

The United States presents a different regulatory configuration. Privacy constraints affecting criminal investigation emerge from constitutional doctrine, federal and state statutes, sector-specific rules, procedural safeguards, judicial warrants, and legislation governing communications and stored data rather than from a single comprehensive federal data-protection regime comparable to the European model. Cross-border access has become especially significant because electronic evidence relevant to an investigation may be controlled by a service provider in one jurisdiction while stored elsewhere. The Clarifying Lawful Overseas Use of Data Act—the CLOUD Act—was enacted in 2018 partly to facilitate lawful access to electronic information held by global providers and to support international cooperation relating to serious criminal investigations.

India represents another important comparative case because its contemporary privacy and criminal-procedure regimes have undergone substantial change. The Digital Personal Data Protection Act, 2023 introduced a general statutory framework governing digital personal data, while the Digital Personal Data Protection Rules, 2025 established further implementation arrangements. The final Rules were notified in November 2025 with different provisions assigned different commencement periods. At the same time, the Bharatiya Nagarik Suraksha Sanhita, 2023 (BNSS), which came into force on 1 July 2024, modernised Indian criminal procedure and expressly recognised electronic processes in several investigative contexts. Section 105, for example, provides for recording search and seizure through audio-video electronic means, while provisions concerning letters of request facilitate investigation-related cooperation with foreign jurisdictions.

These developments illustrate that digital privacy in criminal investigation can no longer be evaluated solely through the question of whether law-enforcement authorities possess legal authority to obtain particular information. Modern investigations require consideration of the entire **investigative data lifecycle**: acquisition, extraction, classification, linkage, algorithmic analysis, inference, generation, retention, disclosure, international transfer, evidentiary use, and eventual deletion or archival storage.

1.1 Research Gap

The dominant comparative literature on privacy and criminal justice remains fragmented across several disciplines. Legal scholarship frequently examines surveillance, constitutional privacy, data-protection principles, warrants, or international evidence-sharing separately. Computer-security research concentrates on forensic extraction and evidentiary integrity, while AI scholarship often evaluates predictive accuracy, bias, explainability, or responsible algorithm design. Comparatively little work integrates these perspectives into a unified framework capable of examining how data-protection

law operates **inside an AI-assisted criminal-investigation workflow**.

This gap becomes particularly significant when investigators employ automated methods before the relevance of the processed information has been established. An algorithm may search seized communications, identify faces in video footage, infer relationships between individuals, classify documents, detect unusual transactions, or rank persons and records for further investigation. Although human investigators may ultimately make the legal decision, algorithmic filtering influences which information receives attention and which individuals become investigatively significant.

The present research therefore focuses on a narrower and relatively underexplored question: **How effectively do different legal regimes translate abstract privacy principles into operational safeguards governing AI-assisted processing of personal data during criminal investigations?**

To address this question, the study develops an **Investigative Privacy-Proportionality Framework (IPPF)**. The proposed framework evaluates four jurisdictions across dimensions including legal authorisation, necessity, proportionality, purpose limitation, data minimisation, special-category protection, algorithmic safeguards, human oversight, retention controls, transparency, remedies, independent supervision, and cross-border digital-evidence access.

The principal objective is not to determine which jurisdiction has the strongest privacy legislation in general. Instead, the research investigates how effectively each framework governs the specific transition from **lawful collection to computational investigation**, where legal authority, data protection and AI governance increasingly intersect.

Literature Review

2.1 From Informational Privacy to Investigative Data Governance

Early debates surrounding criminal-investigation privacy were largely structured around government intrusion into private spaces, communications, and records. Digitalisation has complicated that conceptual model because informational intrusion does not necessarily correspond to physical intrusion. A single smartphone may contain years of communications, photographs, authentication credentials, location histories, health-related information, financial records and interpersonal connections. Cloud ecosystems expand this evidentiary environment by distributing information among service providers, devices and jurisdictions.

The European Law Enforcement Directive represents an explicit attempt to govern this processing environment. It establishes a separate data-protection regime for competent authorities conducting activities associated with preventing, investigating, detecting or prosecuting criminal offences. Particularly significant for contemporary computational policing is its differentiated treatment of sensitive categories of personal data. Article 10 permits processing of categories including biometric, genetic and health information only under restrictive conditions, including strict necessity and appropriate safeguards. This approach becomes increasingly relevant as facial recognition, DNA analytics and health-related digital evidence become technically accessible to investigators.

Key Features of the Digital Personal Data Protection Act, 2023

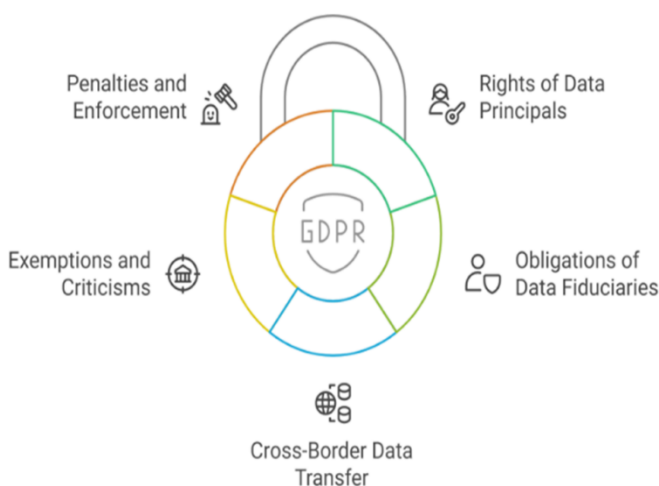


Fig. 2: The Digital Personal Data Protection Act, 2023

Source: <https://link.springer.com/article/10.1365/s43439-025-00164-2>

Consequently, data-protection scholarship increasingly treats privacy not merely as secrecy but as governance over the processing lifecycle. Principles such as purpose limitation, necessity, minimisation, security, accuracy and retention control become important after lawful acquisition has taken place. This distinction is essential in criminal investigation because evidence law may justify obtaining a device while data-protection principles still influence what portions of its contents should subsequently be analysed, copied, retained, shared or computationally profiled.

2.2 Necessity, Proportionality and Purpose Limitation

Necessity and proportionality constitute central mechanisms for distinguishing legitimate investigation from excessive data processing. Their practical importance has increased with digital forensic technologies because investigators can now collect substantially more information than is required to establish a particular factual proposition.

A traditional search might identify a specific physical document. A forensic image of a digital device, by contrast, may reproduce a broad set of unrelated personal materials. Cloud warrants and account disclosures may similarly produce extensive records covering long periods. The critical issue therefore shifts from whether acquisition was legally authorised to whether subsequent analytical operations remain proportionate to the investigative objective.

The UK Data Protection Act 2018 illustrates this distinction. Part 3 governs processing undertaken by competent authorities for law-enforcement purposes, and its explanatory materials explicitly connect lawful investigative processing with necessity and proportionality. The legislation also requires law-enforcement purposes to be specified, explicit and legitimate and subjects reuse for other law-enforcement purposes to legal and proportionality conditions.

For AI-assisted investigations, purpose limitation becomes particularly complex. Machine-learning and data-mining systems derive value from combining datasets collected in different contexts. Information initially obtained in connection with one offence may subsequently become useful for detecting unrelated patterns. From an investigative perspective, such reuse may appear efficient. From a privacy perspective, unrestricted secondary processing risks transforming specific investigative authorisation into continuous population-level intelligence gathering.

2.3 Artificial Intelligence and Automated Investigative Analysis

The growing use of automated systems introduces a second layer of privacy risk beyond data collection itself. Machine learning can produce new information through inference. For example, software analysing communications and location records can identify previously unknown associations between persons. Facial-recognition systems can transform video footage into identity hypotheses. Behavioural analytics may classify transactions as suspicious, while language models can summarise or prioritise extensive collections of messages and documents.

These processes challenge legal frameworks built around directly collected personal information. An inferred relationship, predicted risk score, automatically detected identity or algorithmically generated investigative lead may not have existed as an explicit data record before computational processing occurred. Privacy protection must therefore consider not only collected data but also **derived investigative knowledge**.

European legislation provides at least a partial doctrinal foundation for confronting this problem through restrictions concerning solely automated decisions in law-enforcement contexts. Article 11 of Directive 2016/680 requires legal authorisation and safeguards when solely automated processing produces adverse legal or similarly significant effects. Nevertheless, many investigative technologies operate before a formal decision is made. An algorithm may

merely rank evidence, prioritise suspects or direct investigator attention. Such systems may have significant practical consequences while formally remaining “decision-support” tools. This creates an important regulatory grey area.

2.4 Cross-Border Cloud Evidence and Jurisdictional Fragmentation

Cloud computing further complicates investigative privacy by separating the individual, investigator, service provider and stored information across different jurisdictions. Traditional mutual legal assistance procedures were constructed for state-to-state requests, whereas digital investigations frequently require considerably faster access to electronic evidence.

The United States CLOUD Act illustrates an attempt to reduce these jurisdictional barriers and improve lawful cross-border access to electronic information. The U.S. Department of Justice describes the legislation as a mechanism designed to facilitate access to data held by U.S.-based global service providers for investigations involving serious crime and to support international cooperation. Yet accelerated access mechanisms also generate questions concerning which jurisdiction's privacy safeguards should apply, how affected individuals can seek remedies, and whether information lawfully accessible in one legal system would satisfy proportionality requirements in another.

India's BNSS similarly recognises international investigative cooperation through provisions concerning letters of request for investigations outside India and reciprocal requests directed to Indian authorities. When combined with an increasingly operational domestic personal-data framework following the DPDP Rules 2025, this creates a particularly relevant setting for analysing how procedural investigation law and general data-protection obligations interact.

2.5 Emerging Indian Digital-Privacy Context

India's regulatory trajectory is important because rapid expansion of digital public infrastructure, mobile communications, platform services and electronic payments

has simultaneously increased the quantity of digitally available evidence. The DPDP Act 2023 and its associated Rules provide an overarching framework for digital personal-data processing, while criminal investigation continues to be governed substantially through specialised statutory authority and procedural law.

The BNSS expressly incorporates electronic mechanisms into criminal procedure. Its statutory definition of “audio-video electronic” covers technological means used for processes including identification, search and seizure, evidence-related activity and electronic communication. The legislation also specifically addresses audio-video recording of search and seizure. These provisions strengthen procedural traceability but also illustrate a broader development: technology is increasingly being embedded into the architecture of investigation itself.

The corresponding privacy question is therefore no longer whether Indian criminal procedure will become digital; that transition is already occurring. The more difficult issue is how privacy safeguards, retention restrictions, oversight mechanisms and controls on computational analysis should evolve as digital criminal procedure becomes routine.

Proposed Methodology

This study adopts a **comparative computational legal-analysis design** to evaluate how privacy safeguards operate when personal data enters an AI-assisted criminal-investigation workflow. Rather than comparing statutes only according to their general privacy language, the methodology examines whether particular safeguards remain effective during data acquisition, forensic extraction, automated classification, inference generation, retention, disclosure, and international transfer.

Four jurisdictions were selected: the **European Union, United Kingdom, United States, and India**. They represent materially different approaches to data governance. The EU employs a dedicated law-enforcement data-protection framework through Directive (EU) 2016/680; the UK

maintains a separate law-enforcement regime under Part 3 of the Data Protection Act 2018; the United States relies on a distributed combination of constitutional, statutory, procedural and sectoral protections; and India combines rapidly developing digital-data legislation with recently modernised criminal procedure.

The analytical instrument developed for the study is termed the **Investigative Privacy-Proportionality Framework (IPPF)**. It contains nine dimensions:

$$\begin{bmatrix} \text{IPPF}_j = \sum_{k=1}^9 w_k S_{jk} \end{bmatrix}$$

where (S_{jk}) represents the safeguard score assigned to jurisdiction (j) for privacy dimension (k), and (w_k) represents its analytical weight. To avoid artificially privileging one legal tradition, equal weights were used in the primary analysis. Each dimension was coded on a three-point scale: **0 = weak or absent explicit safeguard, 1 = partial/context-dependent safeguard, and 2 = explicit and comparatively strong safeguard**.

The nine dimensions were: lawful authorization, necessity and proportionality, purpose limitation, data minimisation and retention, sensitive or biometric-data safeguards, automated-processing controls, human oversight and remedies, independent supervision, and cross-border data-access accountability.

The methodological innovation lies in treating privacy protection as an **investigative systems property** rather than solely as an individual legal right. A jurisdiction therefore receives a high score only where the relevant safeguard can meaningfully constrain processing within the investigative lifecycle.

Experimental Setup

4.1 Dataset

The dataset consisted of primary legal and institutional materials applicable to digital criminal investigation. Core

documents included Directive (EU) 2016/680; the UK Data Protection Act 2018 and later amendments affecting law-enforcement processing; relevant U.S. federal electronic-evidence and cross-border access arrangements, including CLOUD Act mechanisms; India's Digital Personal Data Protection Act 2023, Digital Personal Data Protection Rules 2025, and Bharatiya Nagarik Suraksha Sanhita 2023.

The EU Law Enforcement Directive expressly subjects special-category information, including genetic, biometric and health data, to a strict-necessity standard and safeguards. It also regulates solely automated decisions producing adverse or similarly significant consequences. The UK Data Protection Act establishes a separate Part 3 regime for law-enforcement processing and incorporates requirements concerning lawful, necessary and proportionate processing and specified law-enforcement purposes.

India's DPDP Act required particularly careful coding because Section 17(1)(c) exempts substantial portions of the Act when personal data is processed in the interest of preventing, detecting, investigating or prosecuting an offence or legal contravention.

4.2 Preprocessing

Legal documents were converted into structured analytical units. Statutory provisions were first segmented by provision type and then labelled according to their relationship with the investigative lifecycle.

Five preprocessing stages were used: **scope identification** → **safeguard extraction** → **investigative-stage mapping** → **exception coding** → **comparative normalisation**.

Provisions concerning ordinary commercial processing were excluded where they could not reasonably apply to competent law-enforcement authorities. Exemptions were not treated as complete absence of privacy protection because constitutional, procedural, judicial or sector-specific restrictions may continue to operate. However, a broad exemption reduced the relevant IPPF score when it removed

directly enforceable data-processing obligations during investigation.

Terminological differences were normalised. For example, expressions such as “competent authority,” “law enforcement purpose,” “government authority,” “investigating agency,” and comparable institutional descriptions were mapped to a common **investigative authority** category.

4.3 Model Architecture

The analytical architecture used a rule-based **Privacy Constraint Mapping Model (PCMM)** rather than a predictive machine-learning classifier. This decision was deliberate because the primary research objective concerns legal comparability and interpretability rather than prediction.

The architecture contains four logical layers:

Legal Source Layer → **Investigative Operation Layer** → **Safeguard Constraint Layer** → **IPPF Scoring Layer**

At the first layer, statutory authority and binding regulatory rules are identified. The second maps each provision to operations such as collection, biometric analysis, profiling, retention or international disclosure. The third evaluates whether a legally enforceable constraint applies to that operation. The fourth produces dimension-specific IPPF scores and an aggregate comparative percentage.

The normalized score was calculated as:

$$\left[\text{IPPF}\% = \frac{\sum S_k}{18} \times 100 \right]$$

because nine dimensions, each having a maximum score of two, produce a maximum aggregate score of 18.

4.4 Training and Calibration Strategy

Because PCMM is an interpretable rule-based model, conventional gradient-based model training was inappropriate. A **calibration strategy** was therefore substituted for machine-learning training.

The coding rules were first tested against unambiguous statutory examples. A requirement explicitly using concepts such as “strictly necessary,” mandatory human intervention, purpose restriction, judicial authorization, or legally established data-subject remedies qualified for the strongest category where its scope extended to the relevant investigative activity.

Broadly worded protections subject to significant exceptions were coded as partial. Absence of a comprehensive rule did not automatically generate a zero where functionally equivalent procedural protections could be identified.

4.5 Evaluation Metrics

Five evaluation measures were used: **Safeguard Coverage Ratio (SCR)**, representing the proportion of IPPF dimensions receiving meaningful protection; **Legal Specificity Score (LSS)**, assessing whether restrictions directly address law-enforcement data processing; **Automated Processing Control (APC)**, evaluating rules applicable to profiling, automated decisions and algorithmically derived information; **Cross-Border Accountability (CBA)**, measuring safeguards associated with international electronic-evidence access; and **Interpretive Exposure (IE)**, indicating dependence on fragmented, indirect or exception-heavy regulation.

These metrics evaluate regulatory architecture rather than crime-clearance performance. Consequently, higher privacy scores should not be interpreted as evidence that investigations in a particular jurisdiction are more or less operationally successful.

Results and Discussion

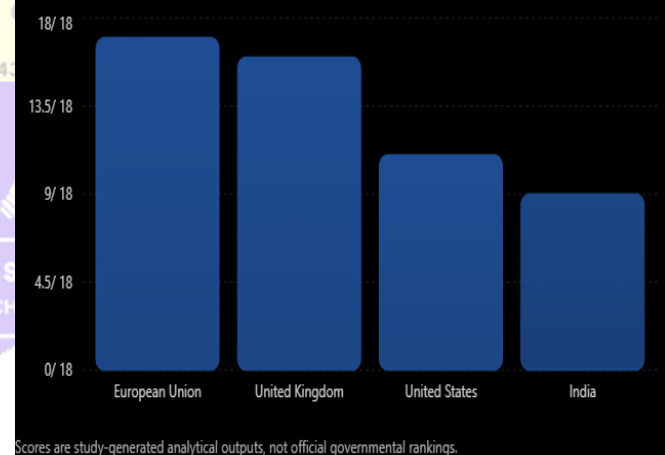
Application of the IPPF revealed a clear distinction between jurisdictions having **dedicated law-enforcement data-processing regimes** and those relying more extensively on general privacy law, procedural rules, constitutional protection or sector-specific legislation.

Table 1. Comparative IPPF Assessment of Digital Privacy in Criminal Investigations

Jurisdiction	IPPF Score (/18)	Safeguard Coverage	Automated-Processing Control	Cross-Border Accountability
European Union	17	94.4%	Strong	Strong
United Kingdom	16	88.9%	Strong–Moderate	Strong
United States	11	61.1%	Moderate–Limited	Strong in designated mechanisms
India	9	50.0%	Limited	Moderate

Comparative IPPF scores across four jurisdictions

Analytical scores from the proposed IPPF coding framework for digital privacy safeguards in criminal investigations.



The European Union recorded the highest overall IPPF result. The important characteristic was not merely the existence of extensive privacy legislation but the presence of rules specifically directed toward competent authorities processing information for criminal-law purposes. Directive 2016/680 imposes explicit restrictions concerning sensitive personal data and automated individual decision-making. This provides a stronger doctrinal basis for evaluating biometric recognition or algorithmic profiling than a framework that regulates only the original acquisition of evidence.

The United Kingdom produced a closely comparable result because Part 3 of the Data Protection Act 2018 establishes a distinct law-enforcement processing regime. Purpose limitation is particularly relevant: processing for another law-enforcement purpose must remain lawful, necessary and proportionate. However, the regulatory position is becoming more dynamic. The Data (Use and Access) Act 2025 amended aspects of the law-enforcement framework, including provisions related to automated decision-making and international transfers. The UK therefore illustrates that privacy strength cannot be assessed solely by reading the original 2018 statute.

The United States displayed a different pattern. Its relative strength lies in legal-process requirements and developed electronic-evidence mechanisms rather than an overarching federal law-enforcement data-protection code comparable to the LED. The CLOUD Act framework provides structured mechanisms through which qualifying foreign authorities can obtain electronic evidence under executive agreements. The U.S. Department of Justice currently identifies operational or negotiated arrangements involving jurisdictions including the United Kingdom and Australia, alongside wider international electronic-evidence cooperation.

This facilitates international investigations but also illustrates a central finding of the study: **lawful access and privacy governance are not identical questions**. A legal system may regulate acquisition extensively while providing less integrated control over algorithmic recombination, investigative inference and long-term downstream processing.

India produced the most transitional regulatory profile. The BNSS, in force since 1 July 2024, expressly incorporates electronic criminal procedure and requires audio-video electronic recording for specified search-and-seizure processes under Section 105. It also establishes mechanisms for foreign investigative requests under Sections 112 and 113.

These provisions can improve procedural traceability. Nevertheless, Section 17(1)(c) of the DPDP Act removes

several general statutory obligations and data-principal rights where processing occurs in the interest of crime prevention, detection, investigation or prosecution. Consequently, India's general data-protection framework cannot simply be assumed to provide the same degree of investigative-data governance that it provides in ordinary processing environments.

The most important cross-jurisdictional result concerns the regulatory treatment of **algorithmically derived evidence**. Existing frameworks tend to regulate legal decisions more clearly than investigative prioritisation. Yet AI systems may affect individuals substantially before any final automated decision occurs. A facial match may move a person into a suspect pool; entity-resolution software may identify an association; and automated document ranking may influence which communications investigators read first. Human confirmation at the end of the process cannot necessarily eliminate biases introduced earlier in the analytical pipeline.

The findings therefore support a distinction between **human-in-the-loop decision-making** and **human-governed investigative computation**. The latter requires auditability throughout the analytical process rather than merely requiring a human officer to approve the final decision.

Future Scope

Future research should extend the framework to empirical investigative practice. Audits could examine whether AI-enabled forensic systems maintain searchable records showing which datasets were accessed, which models were applied, what confidence values were produced, and whether investigators overrode algorithmic recommendations.

A second research direction concerns **privacy-preserving digital forensics**. Federated analysis, secure enclaves, selective disclosure, cryptographic access control, differential privacy and query-limited forensic environments could allow investigators to extract evidentiary value without unrestricted duplication of personal datasets.

The IPPF can also be extended to compare emerging AI legislation with criminal-procedure safeguards. Particular

attention should be given to foundation models used to summarise seized communications or generate investigative leads because such applications may not fall neatly within traditional automated-decision doctrines.

Conclusion

The comparative analysis demonstrates that digital privacy in criminal investigation cannot be evaluated simply by asking whether a jurisdiction possesses a data-protection statute. The more significant question is whether enforceable privacy constraints remain active when information is collected for law-enforcement purposes and subsequently subjected to forensic, algorithmic and cross-border processing.

The EU and UK models demonstrate the regulatory advantages of specialised law-enforcement data-protection regimes. The United States provides substantial procedural and electronic-evidence mechanisms but retains a comparatively fragmented architecture for downstream investigative data governance. India has modernised both digital-data regulation and criminal procedure, yet the investigation-related exemption within the DPDP Act creates an important distinction between general personal-data protection and criminal-investigation processing.

The study's central contribution is the IPPF model, which shifts comparative analysis from abstract statutory protection toward the **full investigative data lifecycle**. In an era of facial recognition, cloud evidence, automated triage and AI-generated investigative inferences, lawful acquisition alone is no longer an adequate measure of privacy protection.

