forensic sciences

MDPI

Article

The Golden Hours of Criminal Investigation: AI-Enhanced Decision Support for Data-Driven, Evidence-Based Policing

Dimitar Rangelov [1,*] and Jaap Knotter [2]

[1] Data Management & Biometrics (DMB), University of Twente, Drienerlolaan 5, 7522 NB Enschede, The Netherlands
[2] Technologies for Criminal Investigations (TCI), Police Academy of the Netherlands and Saxion University of Applied Sciences, Arnhemseweg 348, 7334 AC Apeldoorn, The Netherlands; jaap.knotter@politieacademie.nl
* Correspondence: d.g.rangelov@utwente.nl

Abstract

Background/Objectives: The golden hours, the critical initial phase of each criminal investigation, are crucial for securing evidence, engaging witnesses, and directing investigative efforts, yet this phase often relies heavily on individual expertise and intuition. This paper explores how artificial intelligence (AI) can enhance evidence-based and data-driven policing during these early hours by improving decision-making and investigative quality. **Methods:** Drawing on a source pool of 1589 open-source case records, covering solved and unsolved homicides, suspected homicides, suspicious disappearances, and deaths initially of unclear cause in the Netherlands, compiled from public repositories that aggregate material originating from police files and published court decisions, the study develops a conceptual framework that integrates AI techniques such as automated narrative generation, 3D reconstruction, and digital evidence analysis into early-stage investigations, and illustrates it retrospectively with a purposively selected, coded twenty-case sample; scenario generation in the proposed system follows the twelve elementary scenario components (ESC12). **Results:** By linking empirical data from previous cases with new operational information, that is, information that emerges or is newly discovered during the ongoing investigation, AI is designed to support the narrative approach, accelerate scenario reconstruction, and generate candidate hypotheses and leads for investigator assessment. An author-assigned preliminary evidence-and-readiness assessment rates four AI application domains against five golden-hour criteria, and the retrospective case analysis maps each framework component onto documented investigative problems. **Conclusions:** AI may improve investigative efficiency and consistency and is hypothesised to mitigate the impact of cognitive bias and time pressure; these effects remain to be demonstrated in prospective evaluations. Meanwhile, challenges remain regarding explainability, reliability, and ethical compliance.

Keywords: artificial intelligence; evidence-based policing; data-driven investigations; golden hours; crime scene investigation; narrative approach; 3D reconstruction

Check for updates

Academic Editor: Ricardo Jorge Dinis-Oliveira

Received: 2 September 2026
Revised: 13 October 2026
Accepted: 15 October 2026
Published: 19 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Criminal investigations represent one of the most complex, time-sensitive, and consequential activities within law enforcement. The quality of an investigation's early phase, commonly referred to as the "golden hours", often determines whether a case will be resolved successfully or remain unsolved [1,2]. During these critical initial hours, investigators must rapidly collect physical and digital evidence, engage witnesses, establish

preliminary hypotheses, and decide the direction of the investigation. However, this process is inherently challenging: data arrives in a fragmented, non-linear fashion from diverse sources such as forensic specialists, witnesses, digital systems, and open-source intelligence, and much of it is unverified, incomplete, or potentially unreliable [3,4].

Traditionally, the management of the golden hours has depended heavily on the expertise and intuition of individual (senior) investigating officers (SIOs) and crime scene investigation (CSI) team leaders [5,6]. While experienced investigators can draw on professional judgment, this reliance on individual capability introduces significant variability in investigative quality. Cognitive biases, particularly tunnel vision and confirmation bias, represent well-documented risks in criminal investigations, potentially leading to premature case closure, wrongful convictions, and missed evidence [7–9]. The 'Schiedammer Park murder' case in the Netherlands, for instance, demonstrated how tunnel vision in a police investigating team contributed to a significant miscarriage of justice, prompting nationwide reforms in investigative protocols [10].

The principles of evidence-based policing (EBP), as articulated by Sherman [11], offer a promising framework for addressing these challenges. EBP advocates for policing decisions grounded in rigorous empirical research rather than tradition, assumptions, or convention. Sherman's "triple-T" strategy, targeting, testing, and tracking, provides a systematic approach to deploying police resources based on evidence of effectiveness. However, the application of EBP principles to the dynamic, time-pressured environment of the golden hours remains largely unexplored.

Concurrently, the rise of artificial intelligence (AI) has opened new possibilities for enhancing criminal investigations. AI-driven techniques such as 3D crime scene reconstruction, automated narrative generation, digital evidence triage, and scenario simulation have demonstrated significant potential in forensic contexts [12,13]. Recent advances in Neural Radiance Fields (NeRF) and 3D Gaussian Splatting (3DGS) have enabled rapid, high-fidelity 3D reconstruction from simple video footage, reducing capture time from hours to minutes compared to traditional methods such as laser scanning [14,15]. Similarly, structured reasoning and narrative-based approaches to criminal case analysis [16,17] provide conceptual precedents for the proposal that natural language processing and LLM techniques could assist in generating plausible crime scenarios from available evidence; whether such support in fact accelerates hypothesis formation or mitigates cognitive pitfalls has not been demonstrated empirically and is treated in this paper as a proposed function to be evaluated.

Despite these advances, the integration of AI into the golden hours of criminal investigations faces substantial barriers. AI tools often lack the reliability, explainability, and seamless workflow integration required for time-sensitive, high-stakes decision-making. Their judicial acceptance remains uncertain, particularly when evidence is gathered under the pressures that characterise the earliest phase of an investigation [18,19]. Furthermore, ethical considerations surrounding bias, privacy, accountability, and the potential for automation bias demand careful governance frameworks [20,21].

This paper addresses these challenges by exploring how AI can strengthen evidence-based policing through data-driven methods specifically during the golden hours of crime investigations. The central research question is: What roles can AI technologies play in evidence gathering and decision-making during the golden hours of crime investigations, and what are the associated risks, limitations, validation requirements, and ethical considerations? To answer this question, the study proposes a conceptual framework that integrates AI techniques with established investigative methodologies, drawing on a source pool of 1589 open-source case records collected by the authors from Dutch police records and

published court decisions, complemented by a dataset from the Police of the Netherlands, and insights from investigative practice.

The scientific contribution of this work lies in providing a conceptual framework that links AI technologies with evidence-based policing in the context of early-stage investigations. It expands the theoretical understanding of scenario reconstruction, digital investigations, and AI-supported decision-making, building upon and extending the work of Epskamp-Dudink [3] on scenario thinking, Groenendaal and Helsloot [7] on tunnel vision in Dutch investigations, and our prior research on 3D reconstruction in crime scene investigation [14,15]. The practical contribution offers law enforcement agencies insights into adopting AI tools suited for golden-hour conditions, together with ethical and operational considerations for responsible implementation. This research is part of a broader research line investigating the integration of AI, robotics, and sensing technologies into crime scene investigations, conducted in collaboration between the University of Twente, the Police Academy of the Netherlands, Saxion University of Applied Sciences, and the Technical University of Sofia.

The remainder of this paper is organised as follows: Section 2 reviews related work on golden-hour investigations, evidence-based policing, and AI in forensic contexts. Section 3 describes the methodology, including the conceptual framework and data sources. Section 4 presents the findings on AI-enhanced CSI applications and their potential integration into golden-hour workflows. Section 5 discusses implications, limitations, and future directions. Section 6 concludes with a summary of insights and contributions.

2. Related Work

2.1. The Golden Hours in Crime Investigation

The concept of the golden hours refers to the critical initial phase of a criminal investigation, during which the probability of securing decisive evidence and identifying suspects is highest [1,2]. There is no generally accepted operationalisation of the golden hours in the literature or in operational practice, depending on the source and the type of case, the term may cover the first hours up to roughly the first day of an investigation [22]. In this paper, the golden hours are therefore understood functionally rather than as a fixed number of hours: the period in which the initial influx of unverified information must be triaged and the first investigative decisions are taken. This phase is characterised by the rapid influx of unverified information from multiple sources, including forensic specialists, witness statements, digital evidence specialists, and open-source intelligence analysts. The officer in charge must synthesise this fragmented, non-linear data flow to determine investigative direction, a task that is both cognitively demanding and consequential.

Crime scene investigation itself is grounded in Locard's exchange principle, which holds that every contact between individuals or between an individual and an environment leaves a trace [23]. While Locard's framework was developed in an era focused primarily on chemical, physical, and biological traces, contemporary investigations must also account for a vast and growing body of digital traces that were not available during Locard's professional career. The challenge is not merely one of evidence collection but of evidence integration: connecting disparate data points into coherent narratives that can guide investigative decisions [3].

During the golden hours, crucial decisions must be made under significant external pressure from public prosecutors, media, family members of victims or suspects, and the public at large. The officer in charge must assess the reliability and relevance of incoming information while simultaneously managing organisational and political expectations. In this high-pressure context, objective data and AI-supported tools could help investigators make better-informed decisions [4,6]. Recent forensic research likewise emphasises structured

and standardised procedures for complex crime scene work, such as cold case reviews and large-area searches [24].

2.2. Evidence-Based Policing

Evidence-based policing (EBP) is a paradigm that advocates for the use of the best available research evidence to inform police practices and strategies [11]. Introduced by Sherman in 1998 and further elaborated through the "triple-T" framework of targeting, testing, and tracking, EBP contrasts sharply with traditional policing approaches grounded in intuition, hierarchy, and organisational convention [25]. EBP has gained considerable traction internationally, particularly in the United Kingdom, Australia, and the Netherlands, where professional societies and dedicated research centres have been established [26].

The application of EBP to criminal investigations has been explored through several lenses. Knutsson and Tompson [26] have examined advances in evidence-based policing across multiple domains, while Telep and Lum [27] have investigated the receptivity of police officers to empirical research. However, these contributions have primarily focused on strategic and organisational dimensions rather than on the operational realities of the golden hours. The integration of EBP principles into time-critical investigative workflows remains an open area of inquiry.

Van Dijk, Hoogewoning, and Punch [28] have argued that police leadership must embrace a culture of evidence and reflection to improve investigative quality. This cultural dimension is particularly relevant to golden-hour practices, where individual decision-making under pressure can be significantly improved by embedding evidence-based tools and protocols into standard operating procedures. At the same time, EBP has attracted critical reflection: an exclusive focus on quantitative outcome evidence narrows the questions asked of police research, "what works" is not synonymous with good police work, and in operational practice not all relevant data are available or measurable [29]. The framework proposed in this paper therefore treats evidence-based tools as a complement to, rather than a replacement of, professional judgement.

2.3. Scenario Reconstruction and the Narrative Approach

A significant development in Dutch investigative practice has been the emphasis on scenario thinking and narrative reconstruction. Epskamp-Dudink [3] has provided a comprehensive framework for scenario reconstruction in criminal investigations, combining scientific theories of narrative and anchoring with insights from criminal law and investigative practice. The book "Van watskeburt tot whodunit" offers practical frameworks and tools for scenario-based thinking, emphasising the importance of reconstructing the complete narrative rather than focusing prematurely on individual suspects or evidence items. The link with EBP lies in the shared ambition to base investigative decisions on structured, explicit reasoning rather than on intuition alone: scenario reconstruction disciplines the formation of hypotheses in much the same way that EBP disciplines the choice of interventions.

The narrative approach is rooted in narrative criminology, which emphasises the role of storytelling in how crimes are understood, reconstructed, and ultimately prosecuted [30]. Traditionally, narrative construction has been the responsibility of human investigators, legal professionals, and prosecutors who integrate evidence into coherent accounts. Bringing AI into this process introduces additional epistemological and procedural complexities, as AI-generated narratives operate through probabilistic models rather than experiential judgment [17,31]. It should also be noted that in the practice described by Epskamp-Dudink, narratives and hypotheses are generated by the detectives themselves, based on prior experience and knowledge, which raises the question of whether the right scenarios

are formulated for a given case. The use of datasets and more objective, data-driven scenario generation, in the spirit of EBP, was not addressed in that work. The work of Groenendaal and Helsloot [7] discussed below confirms this concern: procedural safeguards alone did not eliminate cognitive bias, which strengthens the case for data-supported scenario generation.

Groenendaal and Helsloot [7] have examined the tension between precaution and efficacy in major criminal investigations in the Netherlands, finding that oversight policies designed to prevent tunnel vision did not significantly improve efficiency or eliminate cognitive bias from the investigative process. Helsloot's work suggests that criminal investigations could be more efficient and productive if investigators focused on leading evidence rather than attempting to document every detail, a principle that aligns closely with the potential of AI to prioritise and triage evidence during the golden hours [7,8].

2.4. AI in Policing and Forensics

The application of AI in law enforcement has expanded rapidly in recent years, encompassing predictive policing, digital forensics, facial recognition, and crime pattern analysis [32,33]. Systematic reviews further document a growing role for AI in specialised forensic domains such as gunshot wound analysis [34]. In the forensic domain, AI has been applied to crime scene documentation through 3D reconstruction techniques that preserve spatial relationships and visual details that might be missed with traditional photography-based methods [35]. Photogrammetry-based 3D documentation is likewise being extended to a widening range of forensic specialisms [36].

In our previous research, which forms the foundation for the present study, we demonstrated the effectiveness of radiance field-based methods for forensic 3D reconstruction. We showed that camera settings and data capture methods significantly influence the quality of 3D reconstructions using NeRF and Gaussian Splatting, with optimal configurations identified for indoor crime scene scenarios [14,15]. These findings established that 3DGS consistently outperforms NeRF in computational efficiency and noise reduction while maintaining comparable visual fidelity. The present paper builds on this body of work by positioning these 3D reconstruction capabilities within a broader framework for AI-enhanced golden-hour investigations, where rapid scene documentation is essential. Reconstruction data of this kind can also serve as an additional, objective source for narrative generation, supporting data-generated lines of inquiry during the early, unstructured stage of a crime scene investigation and helping to counter tunnel vision.

Furthermore, AI-aided 3D crime scene reconstruction systems have been developed that reduce capture time to 5–15 min per scene using affordable camera equipment, compared to more than 4 h with traditional laser scanning systems such as the FARO X 130 [14,15,20]. These systems enable trajectory simulations, evidence digitalisation, and collaborative remote scene review through integrated web viewers, capabilities that are particularly valuable for golden-hour investigations where multiple agencies and stakeholders must coordinate rapidly.

Beyond scene documentation, AI has been explored for crime scenario generation in separate, as yet unintegrated, demonstrations. Ugai et al. [16] have proposed a logical approach to criminal case investigation that uses structured knowledge graphs to model event sequences, demonstrated on a fictional case, while Hajare and Thalor [37] have explored AI-based crime scene simulation through 3D image processing and semantic segmentation; no single evaluated system yet combines 3D scene representations, time-stamped sensor data, and knowledge graphs in an operational pipeline. However, these systems remain largely experimental, and their integration into operational golden-hour workflows has not been systematically evaluated.

Digital evidence analysis represents another critical area where AI can enhance golden-hour investigations. Bhardwaj and Choudhary [38] have developed AI-based decision support systems for cyber forensic investigations, while Dunsin et al. [12] have provided a comprehensive analysis of AI and machine learning in modern digital forensics. Information fusion techniques, as surveyed by Hu et al. [32], offer methods for integrating diverse evidence streams into coherent analytical frameworks. These capabilities are directly relevant to the golden hours, where investigators must rapidly triage and integrate digital evidence from mobile devices, social media, surveillance systems, and communication networks.

2.5. Ethical Considerations and Challenges

The deployment of AI in criminal investigations raises significant ethical concerns that must be addressed before operational integration can proceed responsibly. Bias in AI systems is a well-documented issue in law enforcement contexts, with predictive policing algorithms shown to reinforce systemic disparities by reproducing patterns found in historically biased datasets [39,40]. Critical analyses of predictive policing have further shown how feedback loops between police deployment and recorded crime, and the use of "dirty data" generated during periods of unlawful or biased practice, can entrench discriminatory patterns and undermine the evidential value of historical police data [41–44]. Feedback-loop effects of this kind have been demonstrated formally [45], legal scholarship has questioned the assumptions, evaluation, and accountability of algorithmic prediction in policing [46,47], and empirical work has shown that widely used algorithmic risk predictions can fail to outperform untrained human judgement [48]. In the context of scenario generation, these risks take on new dimensions: if an AI model has been trained on unbalanced data, it may suggest scenarios that reflect those biases even in the absence of supporting evidence [31].

Automation bias, the tendency for human operators to over-rely on automated systems, is a well-documented phenomenon in aviation and clinical decision support [49–52] and is known to be amplified in time-critical decision environments [53]. It therefore presents a particular risk in golden-hour investigations, where time pressure may increase the temptation to accept AI-generated recommendations without adequate critical evaluation [54,55]. The narrative quality of AI-generated scenarios makes them cognitively compelling, as people tend to prefer coherent explanations over disjointed ones, even when the latter are more accurate. This narrative framing effect can strongly influence how investigators perceive a case, potentially leading to premature hypothesis fixation.

Transparency, explainability, and reliability are essential requirements for AI systems deployed in criminal justice contexts. Without explainable outputs, AI-generated scenarios may lack the credibility needed for judicial proceedings or even within investigative teams [18,19]. The EU legal framework imposes additional requirements on AI systems used in law enforcement. Processing of personal data by competent authorities for the prevention, investigation, detection, or prosecution of criminal offences falls outside the material scope of the GDPR (Article 2(2)(d)) and is instead governed by the Law Enforcement Directive (EU) 2016/680 and its national implementations [56]. The EU AI Act (Regulation (EU) 2024/1689) classifies certain law-enforcement uses of AI, including systems intended to evaluate the reliability of evidence, as high-risk; the applicable obligations depend on the intended use of each component rather than on the technology as such [57]. For scholarly analysis of this framework, see [58].

As part of this broader research line, we have previously proposed an ethics-by-design framework for AI-generated crime scene scenarios [31], incorporating human-in-the-loop control, transparency, auditability, and participatory system design as core principles. That work, together with our analysis of the impacts, benefits, and limitations of 3D reconstruction in crime scene investigation [20], provides the ethical foundation upon which the present

study builds. The current paper extends these insights by examining how ethical safeguards can be operationalised specifically within the time-pressured context of the golden hours.

2.6. Research Gap

While significant progress has been made in individual domains, evidence-based policing, AI in forensics, scenario reconstruction, and ethical AI governance, a comprehensive framework connecting these elements specifically within the golden-hour context is lacking. Existing literature addresses these topics separately, without systematically examining how AI can enhance evidence-based decision-making during the critical initial phase of criminal investigations. This paper aims to fill this gap by proposing an integrated conceptual framework and evaluating its potential through analysis of real investigative data.

The distinction from adjacent work can be made explicit. Broad reviews of AI in law enforcement and digital forensics [12,32,33] map available technologies but are not tied to a specific operational phase. Predictive policing frameworks [41,44,46] operate at the strategic level of resource allocation, before any crime has occurred, while digital forensic process models [12,38] concern laboratory-phase analysis after evidence has been secured. The present framework occupies the space between these two: operational decision support during the first hours of an active investigation. To our knowledge, based on the literature reviewed in this paper, it is the first to combine rapid scene documentation, automated scenario generation, digital evidence triage, and OSINT analysis into a single golden-hour decision-support cycle, to assess these components against criteria specific to time-pressured investigative work, and to ground this assessment in a source pool of case records collected by the authors and a purposively selected sample coded by them.

3. Methodology

3.1. Research Approach

This is a conceptual and exploratory study. It combines qualitative conceptual framework development with a descriptive analysis of a collection of documented criminal cases and an illustrative retrospective case mapping; it does not provide an empirical validation of investigative effectiveness. The research design integrates three complementary components: a structured (narrative) literature review of AI technologies and evidence-based policing practices, a conceptual framework linking these domains within the golden-hour context, and the construction of a source pool of 1589 open-source case records originating from police and judicial records, from which a purposively selected sample of twenty cases was coded and described, complemented by a dataset from the Police of the Netherlands.

Table 1 summarises, for each component of the design, the data used, the analytical procedure, where its results appear, and the type of inference it supports; the components are described in Sections 3.2–3.5.

The conceptual approach is deliberately chosen because the integration of AI into golden-hour practices is at an early stage, and a purely empirical approach would be premature without first establishing the theoretical foundations and practical requirements. This aligns with the multi-layered research methodology adopted across the broader research line, which integrates technical experimentation, comparative evaluation, applied analysis, and conceptual reflection.

Table 1. Overview of the research design: components, data, procedures, and the inference each component supports.

Component	Data/Input	Procedure	Output Location	Inference Supported
Structured literature review	Publications 2018–2025 from the databases in Section 3.2	Thematic narrative synthesis	Sections 2 and 4.3	Conceptual grounding; no exhaustive coverage claim
Open-source case corpus ($n = 1589$)	Public case repositories (Section 3.3)	Source-pool construction: cataloguing, de-duplication, and identification of documented investigative problems	Section 4.5	Source-pool coverage and documented problems; no corpus-level findings or prevalence estimates
Coded sample ($n = 20$)	Purposive selection from the corpus	Structured coding by the authors (Table S1)	Table 2, Section 4.5	Illustrative description of the sample only
Retrospective mapping ($n = 5$)	Five cases from the coded sample	Counterfactual mapping of framework components to documented issues	Section 4.6	Illustrative plausibility; no effect or outcome claims
Police archive dataset	Anonymised cold-case and older solved-case files from the police archive (Section 3.3)	Qualitative file review; meetings with police personnel	Sections 3.3 and 4.5	Contextual corroboration; no systematic findings
Practitioner consultation	Informal expert feedback from twelve practitioners (Section 3.4)	Formative consultation	Sections 3.4 and 4.1	Consultative input; not empirical evidence

Table 2. Two example case records from the open-source case collection.

Field	Example Record 1	Example Record 2
Case	C-01 (Etten-Leur family homicide)	C-02 (Bemmelse Waard homicide)
Year	2020	2003
Status	Solved	Solved
Location	Etten-Leur	Bemmelse Waard
Victim(s)	H. K. (64), G. K. (31), A. K. (6), D. K. (2)	M. B. (16)
Suspect/perpetrator	O. K. (33)	F. Ö. (19), G. M. (18), G. P. (17)
Cause of death	Strangulation	Strangulation, burning
Narrative summary	Four family members were found dead in their home on 29 March 2020. The father initially claimed to have been abducted, but camera footage and telephone data disproved his account, and he admitted strangling his mother.	Strangled by three acquaintances in 2003; the body was burned and found in the Bemmelse Waard on 17 November 2003. The perpetrators confessed and received sentences of 3, 8, and 11 years.

3.2. Literature Review Method

The academic literature was gathered from IEEE Xplore, Google Scholar, Scopus, and domain-specific repositories, focusing on publications from the last seven years (2018–2025) unless historical context required older sources. Key search terms included "evidence-based policing", "golden hours criminal investigation", "AI crime scene investigation", "3D reconstruction forensics", "narrative approach criminal investigation", and "ethical AI policing". The review is a structured narrative review rather than a systematic review: searches were run iteratively during the research programme, with the final search performed in January 2026, without a registered protocol, formal screening flow, duplicate logging, or quality

appraisal; sources were selected on relevance to the golden-hour context, methodological transparency, and recency, with preference for peer-reviewed work. The databases and search terms are reported for transparency, the synthesis is thematic, and the novelty claims in this paper are made relative to this review.

3.3. Case Corpus and Coded Sample

The study draws on multiple data sources. A substantial part of the case data was collected by the authors from open-source repositories documenting solved and cold cases in the Netherlands, including Moordzaken (moordzaken.com), Moordatlas (moordatlas.nl), Uitspraken.nl, and Rechtspraak.nl, the official publication platform of the Dutch judiciary. Although these repositories are publicly accessible, the information they contain largely originates from police investigations and judicial proceedings; Rechtspraak.nl publishes the official decisions of the courts themselves, whereas the other repositories are secondary aggregators of such official material and contemporaneous reporting. These open sources provide case descriptions, investigative timelines, and judicial outcomes that make the underlying material available in a transparent manner; in this study they were used to construct the source pool from which the illustrative sample was drawn, and no corpus-wide pattern analysis is reported.

In total, 1589 solved and unsolved (cold case) cases were collected from these open sources; the corpus comprises murders and manslaughters as well as suspicious disappearances and deaths initially of unclear cause in which a crime could not be excluded, and these categories are treated together only for cataloguing purposes. The corpus size corresponds to the case records catalogued in the consulted repositories at the time of collection, with Moordatlas.nl as the largest single source; cases documented in more than one repository were identified by case name, location, and year, and counted once. The source pool was compiled between 2024 and 2026. The conceptual framework and the coding categories were developed before the illustrative cases were selected; the twenty cases were subsequently chosen purposively to illustrate the framework components, so the sample can illustrate, but cannot independently establish, the comprehensiveness or utility of the framework. Screening of the source pool was confined to case identification and de-duplication; no corpus-level coding was performed. For unsolved cases the ground truth is unknown. They are therefore used for descriptive illustration only, while any future accuracy-oriented validation will have to rely on solved cases, for which the outcome is known. From this collection, an illustrative sample of twenty cases, comprising eighteen solved cases and two cold cases and spanning the period 1975–2025, was documented in a structured overview. The sample is not intended to be statistically representative: the cases were selected by the authors to reflect the diversity of case types, evidence constellations, and investigative trajectories present in the corpus. Cases were eligible for inclusion if they concerned a homicide, suspected homicide, or suspicious disappearance in the Netherlands and the publicly available documentation was complete enough to reconstruct the investigative timeline and outcome; cases with insufficient documentation were excluded. For each included case, the variables described below were extracted. Coding was performed by the first author, and the second author independently reviewed all twenty coded cases; initial coding differences arose in three of the twenty cases, concerned the boundaries between coded categories rather than the documented facts, and were resolved by discussion against the cited source material. The sample serves to illustrate the structure of the data and the type of coding applied to the selected cases. Table 2 shows two example records, victims and perpetrators are identified by initials, and full details remain traceable through the cited open sources.

Each case record follows a fixed coding scheme consisting of a case identifier and title, status (solved or cold case), year, location, victim(s) and age(s), suspect(s) or perpetrator(s), cause of death, and a narrative summary of the investigation and its outcome. This uniform structure supports cataloguing of the collection; in this study, descriptive counts, such as the distribution of cases per decade, per cause of death, and per judicial outcome, were produced only for the purposively selected twenty-case sample (Section 4.5 and Supplementary Table S1), and no corpus-level statistics are reported.

In addition to the open-source corpus, the study drew on a complementary dataset from the Police of the Netherlands. This material was obtained directly from the police archive, comprises approximately sixty case files, and consists predominantly of cold cases, together with a small number of older solved cases; under the applicable archive rules, solved cases become available for research purposes only ten years after closure, so recent solved cases could not be included. The material was examined through qualitative file review, complemented by meetings with police personnel; no quantitative analysis was performed on it. The dataset is anonymised, cannot be shared or independently replicated, and its composition reflects what was made available under the data-sharing agreement rather than a random sample of Dutch investigations. For these reasons, it is used only as contextual input consistent with the literature and the illustrative cases and as input to the practical requirements of the framework; no conclusion in this paper rests solely on it.

3.4. Practitioner Consultation

Insights from twelve investigative practitioners, including experienced senior investigating officers and forensic specialists, informed the practical requirements of the conceptual framework. Their input was collected through informal interviews, workshops, and experimental showcases, in which intermediate results were demonstrated and feedback was gathered. This input constituted formative consultation rather than research data: it was not collected under a structured interview protocol and was not systematically recorded or analysed, and it is therefore not presented as empirical evidence in this paper. It informed the practical requirements of the framework and the interpretation of the case material. These consultation activities fall under the ethical approval of the broader PhD project (University of Twente, CIS Ethics Committee, application no. 240519), which covers interviews, focus groups, and workshops with practitioners.

3.5. Analytical Framework

The analytical framework evaluates AI technologies for golden-hour integration across five criteria: (1) speed, the ability to deliver actionable results within the time constraints of the golden hours; (2) accuracy, the reliability and fidelity of AI-generated outputs; (3) explainability, the degree to which AI outputs can be understood and justified by investigators and in judicial proceedings; (4) legal/governance readiness, the regulatory and ethical implementation burden of the defined use case under the applicable legal framework; and (5) adoption potential, the feasibility of integrating AI tools into existing investigative workflows under organisational and technical constraints. These criteria are applied systematically to each AI application domain examined in the study, and the outcome of this assessment is summarised in the author-assigned evidence-and-readiness assessment presented in Section 4.3 (Tables 3 and 4). Figure 1 gives a conceptual overview of the proposed framework.

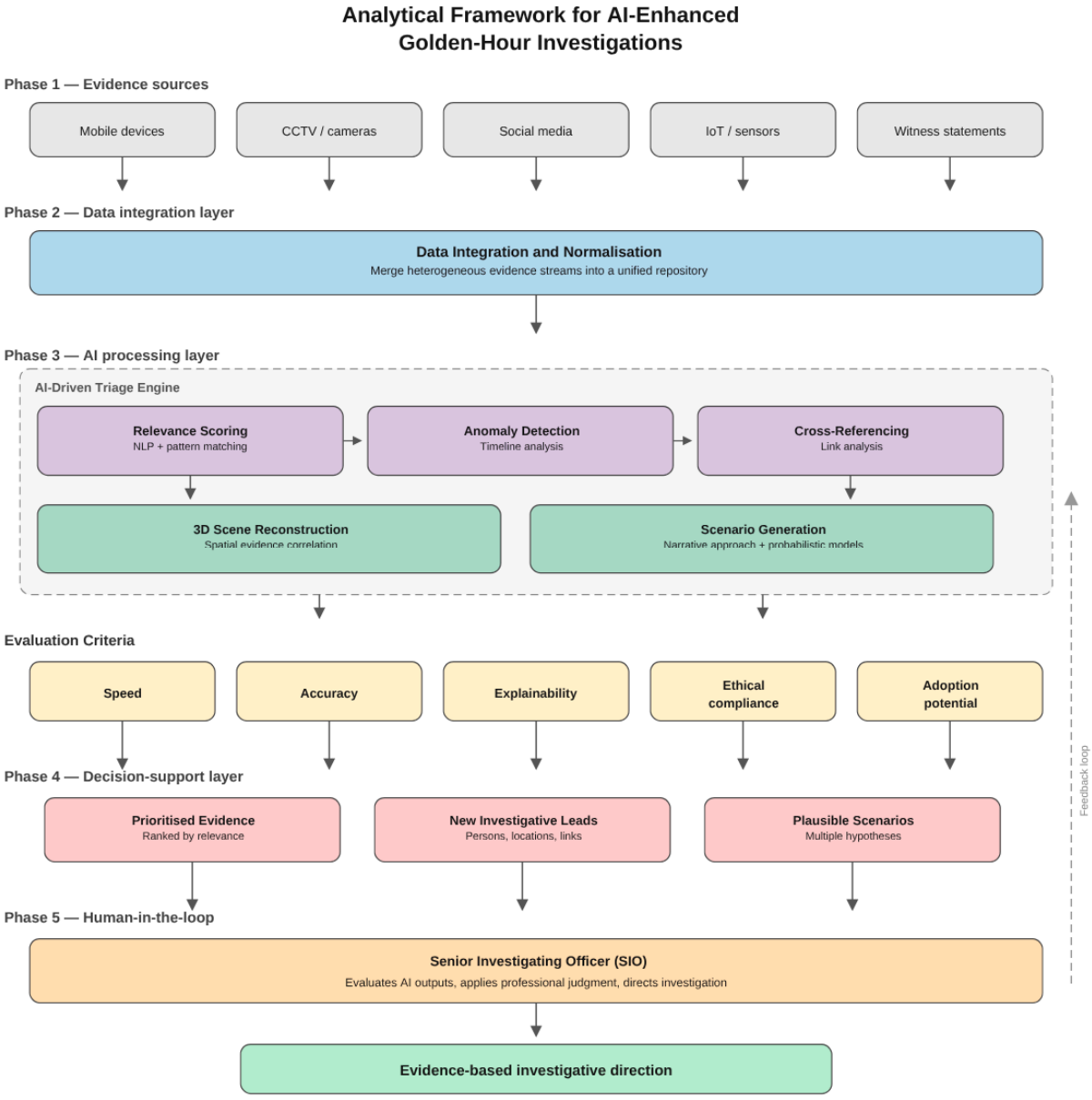


Figure 1. Analytical framework for evaluating AI technologies for golden-hour integration, showing evidence sources, AI processing modules, five evaluation criteria, and the decision-support layer with human-in-the-loop feedback (Conceptual diagram of the proposed framework).

4. Results and Findings

4.1. Conceptual Framework for AI-Enhanced Golden-Hour Investigations

Based on the literature review, practitioner insights (Section 3.4), and analysis of the case data, we propose a conceptual framework for integrating AI into golden-hour investigations. The framework comprises three interconnected layers: (1) the data integration layer, which aggregates information from crime scene documentation, digital evidence sources, witness statements, and historical case databases; (2) the AI processing layer, which applies techniques such as 3D reconstruction, automated narrative generation, evidence triage, and pattern recognition to transform raw data into actionable intelligence; and (3) the decision-support layer, which presents AI-generated insights to investigators in an explainable, contextualised format that supports, but does not replace, human judgment. Hypothesis generation in the AI processing layer follows the same construction as the twelve elementary scenario components (ESC12) introduced by De Kock [17], in which

every conceivable narrative is composed from a fixed set of components covering, among others, the actors, motivation, means, location, time, and modus operandi; the mechanism by which our system generates scenarios is described in Section 4.2.

The framework is designed around the principle that AI should function as an investigation aid rather than as a source of unquestionable truth. This principle aligns with the evidence-based policing paradigm, which positions empirical evidence as a supplement to, not a replacement for, professional experience [11]. The framework also incorporates the scenario reconstruction methodology advocated by Epskamp-Dudink [3], using AI to support the generation and evaluation of multiple plausible scenarios rather than converging prematurely on a single hypothesis. For example, when a partner is found dead in a shared home, the system maintains, next to the domestic-violence scenario that investigators tend to consider first, alternative scenarios such as an intruder, an accident, or suicide, and attaches to each scenario the evidence that supports or contradicts it, so that discarding a scenario becomes an explicit, evidence-based decision.

4.2. AI-Enhanced CSI Applications for the Golden Hours

3D Crime Scene Reconstruction. Rapid 3D reconstruction represents one of the most immediately applicable AI technologies for golden-hour investigations. Our previous research has demonstrated that Gaussian Splatting-based systems can produce high-fidelity 3D reconstructions suitable for documentation and review from video footage captured in 5–15 min, with processing completed within 1–2 h [14,15]; verification and quality review of the resulting model require additional, case-dependent time that has not been separately benchmarked, and capture, processing, and verification times should be reported separately in future evaluations. This represents a substantial reduction compared to traditional methods, which may require more than 4 h for capture alone. The reconstructed 3D environments enable investigators to revisit the crime scene virtually, examine spatial relationships from multiple perspectives, and identify evidence that might have been overlooked during the initial physical examination.

In golden-hour contexts, the value of 3D reconstruction is not limited to recording the scene: the reconstruction itself becomes a working environment in which the investigation can continue. Interactive 3D models support collaborative analysis, allowing multiple investigators and specialists to explore the scene simultaneously from different locations through web-based viewers. This capability is particularly valuable in complex cases involving multiple crime scenes or requiring coordination between agencies. Furthermore, 3D reconstructions provide a repeatable, reviewable record of the scene that reduces dependence on any single observer's documentation choices, an advantage over traditional photography and note-taking. The record is not fully independent of observer bias, however: capture choices, occlusion, reconstruction artefacts, and subsequent interpretation all influence what is represented, and visual realism should not be equated with geometric accuracy or with suitability for a particular forensic inference. Likewise, examining a scenario within a reconstructed environment cannot, without validated physical and temporal constraints, establish that the scenario could in fact have occurred; the demonstrations in this paper are intended as documentation and review aids within those limits.

Beyond spatial documentation, AI-driven object detection applied to 3D reconstructions enables automated identification and classification of evidence items within the scene. This capability supports rapid evidence cataloguing and can highlight objects that might be overlooked during manual inspection, particularly in cluttered or complex environments. Within this paper, these reconstruction and detection capabilities are above all an additional data source: the intention is to combine them with the open-source and operational case

data described in Section 3, so that hypotheses and scenarios can be generated more easily and more reliably from multiple sources, textual and visual, together.

At present, the system generates scenarios as structured text only. In future work, we aim to combine the generated scenarios with the 3D reconstructions, so that plausible narratives can be examined directly within the reconstructed, and potentially VR-based, scene, with all supporting and contradicting evidence integrated into the narrative. This would allow investigators, prosecutors, and judges to verify for themselves whether a narrative could have taken place (Figure 2).

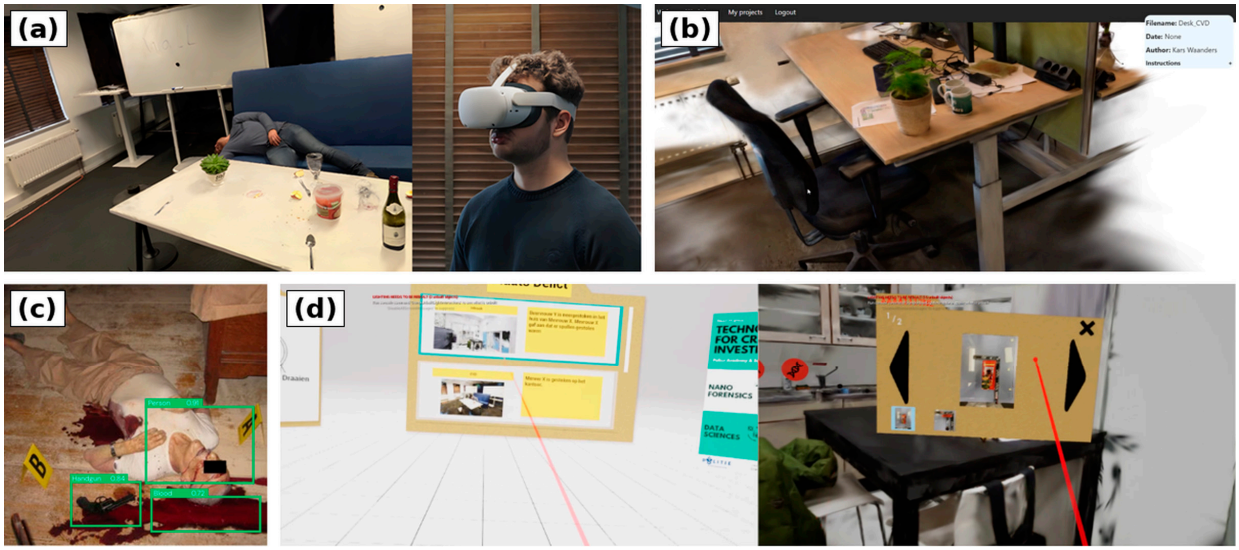


Figure 2. AI-enhanced 3D crime scene documentation and analysis: (**a**) AI-aided 3D reconstruction from video footage using Gaussian Splatting, with VR-based review of the reconstructed scene; (**b**) integrated web viewer for remote collaborative exploration with contextual evidence display; (**c**) AI-based object detection applied to a reconstructed scene; (**d**) interactive VR environment supporting both investigation and training (Illustrative material from the authors' prior experimental work).

As part of the integrated sensing approach, experimental setups involving semi-autonomous robotic platforms have been developed to combine 3D scene reconstruction with real-time chemical and environmental analysis. These systems, equipped with advanced sensor technologies, can enter crime scenes to capture multi-modal data while minimising contamination and operator risk. By combining visible data with non-visible data, such as chemical and audio signatures, both in the reconstruction of crime scenes and in the generation of possible narratives, this combination is a distinctive feature of our prototype; a scoped comparison with related systems remains to be performed.

The applicability of these reconstruction techniques is not limited to a single scene type. Figure 3 shows 3DGS reconstructions of four mock-up scenes created within the research line, covering a wildlife crime scenario, a traffic fatality, an indoor ambiguous-death scene, and scattered remains in outdoor vegetation. This diversity mirrors the range of case types present in the case corpus described in Section 3.3 and indicates that rapid scene documentation can support golden-hour investigations across markedly different investigative contexts, including outdoor scenes where weather conditions can rapidly degrade traces (Figure 4).

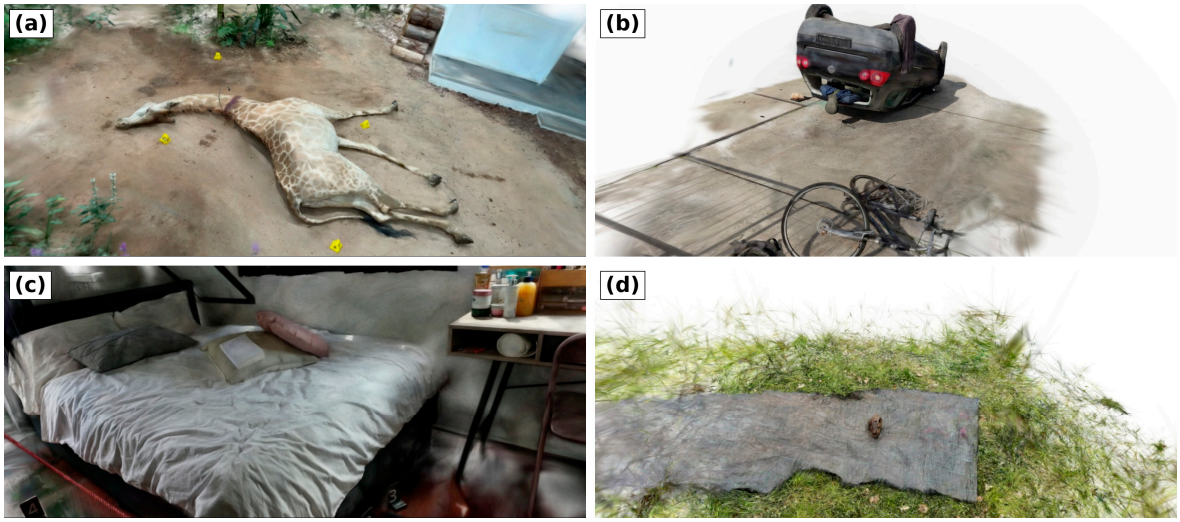


Figure 3. 3DGS reconstructions of mock-up scenes across case types: (**a**) wildlife crime; (**b**) traffic fatality; (**c**) indoor ambiguous-death scene; (**d**) outdoor scattered remains. All scenes are training mock-ups; no real victims are depicted (Prior experimental material).



Figure 4. Experimental setup with a semi-autonomous robot for chemical air analysis and 3D crime scene reconstruction (Experimental prototype from the research programme; illustrative).

Automated Narrative Generation. AI-driven narrative generation systems are proposed to analyse the available evidence, including 3D scene data, digital traces, forensic reports, and witness statements, in order to generate plausible scenarios of what may have occurred. The conceptual precedents are the logical, knowledge-graph-based case reasoning of Ugai et al. [16], demonstrated on a fictional case, and the narrative theory and ESC12 scenario components of De Kock [17]; neither constitutes an empirical demonstration of LLM-based scenario generation in operational investigations, and the operational effects claimed for such systems remain proposed functions to be evaluated. Such systems can operate through probabilistic models, rule-based engines, or neural networks to produce internally consistent and data-supported possibilities. The goal is not to identify a single "true" version of events but to generate a range of alternative narratives that investigators can evaluate and refine as additional evidence becomes available. Generated narratives thus function as research lines in a criminal investigation, and newly collected data can be fed back into the database, so that the set of narratives is continuously updated.

In the golden-hour context, automated narrative generation can accelerate the process of hypothesis formation by systematically combining available data points into coherent scenarios. This capability is particularly valuable when investigators are dealing with the fragmented, non-linear information flows that characterise the early phase of an investigation [3,4]. By presenting multiple scenarios simultaneously, AI can help investigators avoid

premature hypothesis fixation and maintain a broader investigative perspective, directly addressing the tunnel vision risks identified by Groenendaal and Helsloot [7] and Elaad [9].

Within our research line, this concept has been operationalised in our scenario-generation system. The system is implemented in Python 3.11 as a custom pipeline built on open-weight transformer models; in the current prototype, entity and relation extraction and hypothesis generation are performed by a locally hosted, instruction-tuned large language model (Llama 3.3 70B Instruct), steered by task-specific prompt templates and rule-based constraints rather than by fine-tuning. Retrieval selects entities and relations from the structured repository by matching the task specification against the stored graph, and the optional reference documents are formatting exemplars, namely the police-report format and the ESC12 component structure, so that each generated hypothesis is composed over the twelve elementary scenario components and expressed as a police-style report. The case source pool described in Section 3.3 was not used for model training or for retrieval; it served only for framework development and for the selection of demonstration material. Multimodal input handling is implemented for textual and image-derived inputs, whereas audio and sensor-based inputs are proposed capabilities of the architecture that have not yet been exercised end to end. The system remains a research prototype, and a full technical specification and validation protocol will accompany its dedicated evaluation study. The system is designed to accept multimodal crime scene data, textual, visual, audio, and sensor-based, together with police reports (of these, textual and image-derived inputs are implemented in the current prototype, as specified below), and generates up to three plausible alternative hypotheses about the incident in the format of a police report. The limit of three hypotheses is a design choice of the current prototype intended to keep outputs reviewable under time pressure; it is configurable, and the risk that a relevant alternative falls outside the presented set is treated as a failure mode in Section 5.3. As illustrated in Figure 5, incoming multimodal data (1) are first cleaned and pre-processed per modality (2) and transformed into a unified data representation (3). Entities of interest, including persons, objects, locations, events, and temporal markers, and the relations between them are then extracted (4) and stored in a structured, machine-interpretable data repository (5). Given a task specification, the system retrieves the relevant entities and relations (6), optionally combined with reference documents that determine the structure or format of the intended output (7). A reasoning module then generates up to three hypotheses describing plausible crime scenarios (8), which are presented to the user (9). The modular, client-oriented architecture is designed to integrate readily with existing information environments and investigative workflows. As a synthetic end-to-end illustration unrelated to any real case: given a fictitious report that a resident has been found deceased at the foot of an indoor staircase, together with scene photographs, the pipeline extracts entities (the deceased, the staircase, a damp floor mat, an unlatched rear door, a missing wallet) and their spatial and temporal relations; the knowledge graph links each element to its source; and the reasoning module returns three ESC12-structured hypotheses, an accidental fall (anchored to the damp mat; graded F1–F2), a medical event preceding a fall (compatible with but not stated in the sources; F3 assumptions flagged), and a burglary-related confrontation (anchored to the unlatched door and missing wallet; mixed F1–F3), each presented with per-component provenance grades and explicit information gaps, such as the undocumented time of death, for human review.

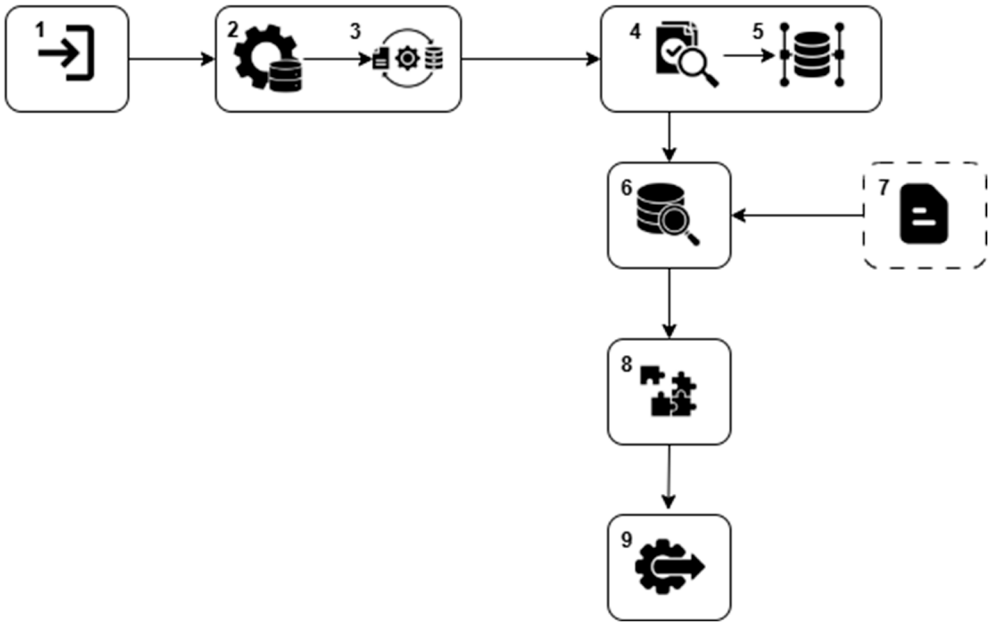


Figure 5. Architecture and workflow of the scenario-generation system: (1) multimodal data input; (2) pre-processing; (3) transformation into a unified data representation; (4) entity and relation extraction; (5) structured data repository; (6) retrieval of relevant data; (7) optional reference documents; (8) reasoning and hypothesis generation; (9) output presentation (Proposed architecture of the prototype).

The system is delivered through an investigator-facing application that guides the user through case setup, review of automatically extracted entities, events, and relations, an interactive case knowledge graph in which every node and edge is linked to its supporting source evidence (Figure 6), and the identification of information gaps. Generated hypotheses are presented with component-level evidential grading, ranging from established facts (F0) to required assumptions (F3), so that investigators can see for every element of a scenario whether it rests on source evidence or on an assumption (Figure 7). The four levels are defined as follows: F0 denotes a statement explicitly documented in a linked source item; F1 a minor inference directly supported by linked sources; F2 a plausible interpretation consistent with, but not explicitly stated in, the sources; and F3 a required assumption without direct source support. The grades record source provenance only: they indicate where a statement comes from, not that it is true. Source reliability, corroboration across independent sources, and the inferential status of a statement are distinct properties that the grading does not collapse; a witness allegation or an inaccurate secondary report may be graded F0 while remaining unverified, and reviewers are instructed to read F0 as documented, not as established. Because consistency with the sources does not by itself provide affirmative support, F2 marks compatibility only, whereas affirmative support requires the direct source linkage of F0–F1; reviewers distinguish the two by asking whether a linked source would be contradicted if the component were false. Disagreements between the two reviewers are resolved by discussion to consensus, and a component that cannot be tied to a source is retained only as an explicitly flagged F3 assumption, so that unsupported assumptions cannot acquire apparent evidential authority through the grading system. Grades are proposed automatically from the source linkage for part of the inputs, are assigned by a human reviewer in all cases, and the assignments are checked by a second person. This explicit provenance grading is intended to mitigate the explainability and automation-bias concerns discussed in Section 2.5 by supporting critical review of every generated component, and the workflow concludes with scenario comparison, a validation report, an investigator review step, and export for documentation. The demonstration data

shown in Figures 6 and 7 are derived from a publicly documented cold case and are presented in anonymised form; the figures are curated demonstrations prepared for publication: they reproduce the application views, but the displayed content was selected, anonymised, and partly re-rendered by the authors for readability and privacy rather than exported unedited from a live session. The hypotheses visible in the interface are demonstrations, not documented conclusions about the underlying case; generated content is visually separated from documented source material by the provenance grades, and the plausibility labels visible in Figure 7 are author-assigned illustrative annotations, not calibrated model outputs, and carry no validated probabilistic or comparative meaning. The demonstration case was transformed before publication: names and initials were replaced by neutral labels, and specific locations, dates, ages, and distinctive circumstances were removed or generalised in the displayed views, so that the combination of displayed attributes does not identify the individuals involved; the residual re-identification risk was assessed by the authors against the publicly documented features of the case before inclusion. Generated allegations and hypotheses are labelled as generated and are kept visually distinct from documented case information, and a fully synthetic demonstration is planned for the dedicated technical publication as a clearer and safer illustration.

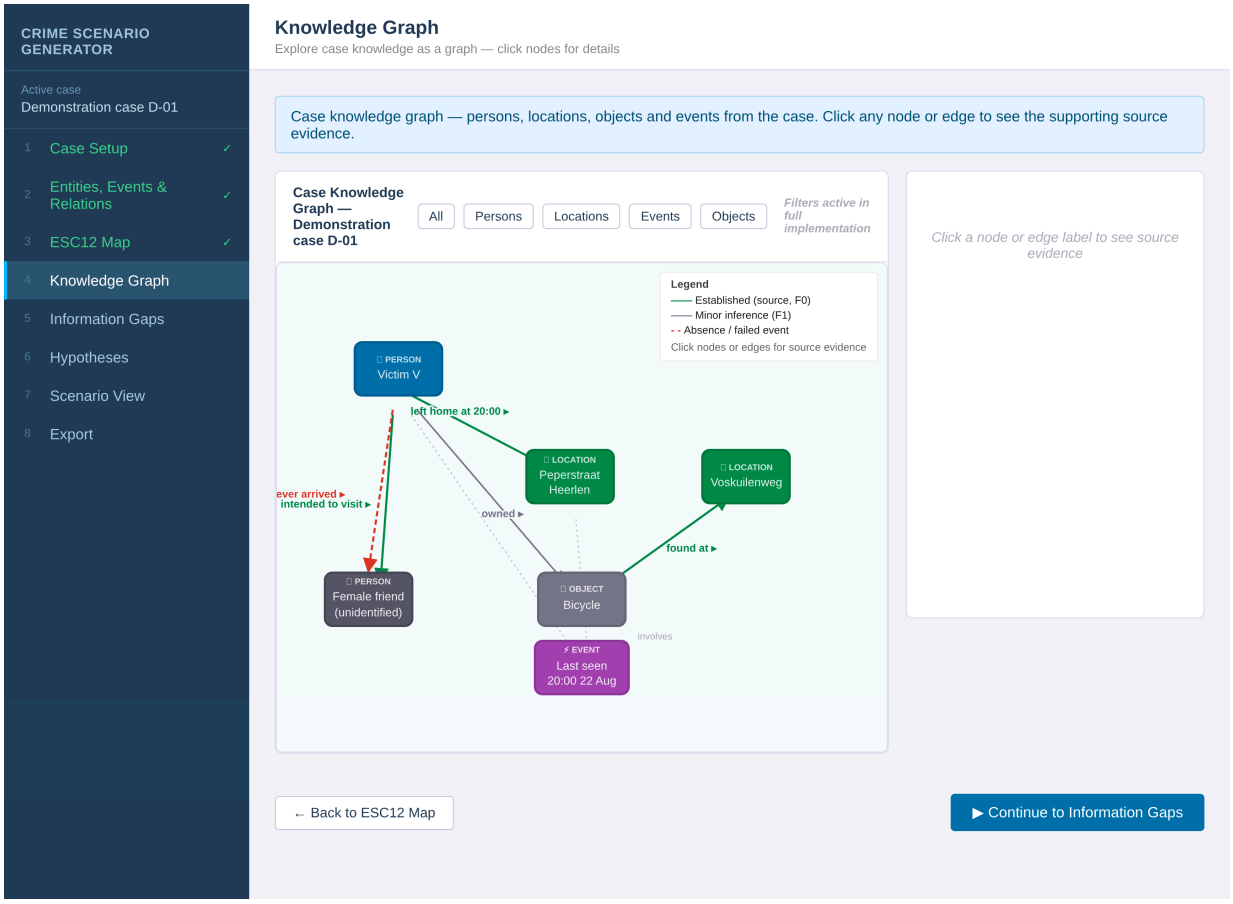


Figure 6. The investigator-facing application: interactive case knowledge graph with evidence-linked nodes and edges, illustrated with anonymised demonstration data derived from a publicly documented Dutch cold case (Interface of the working prototype).

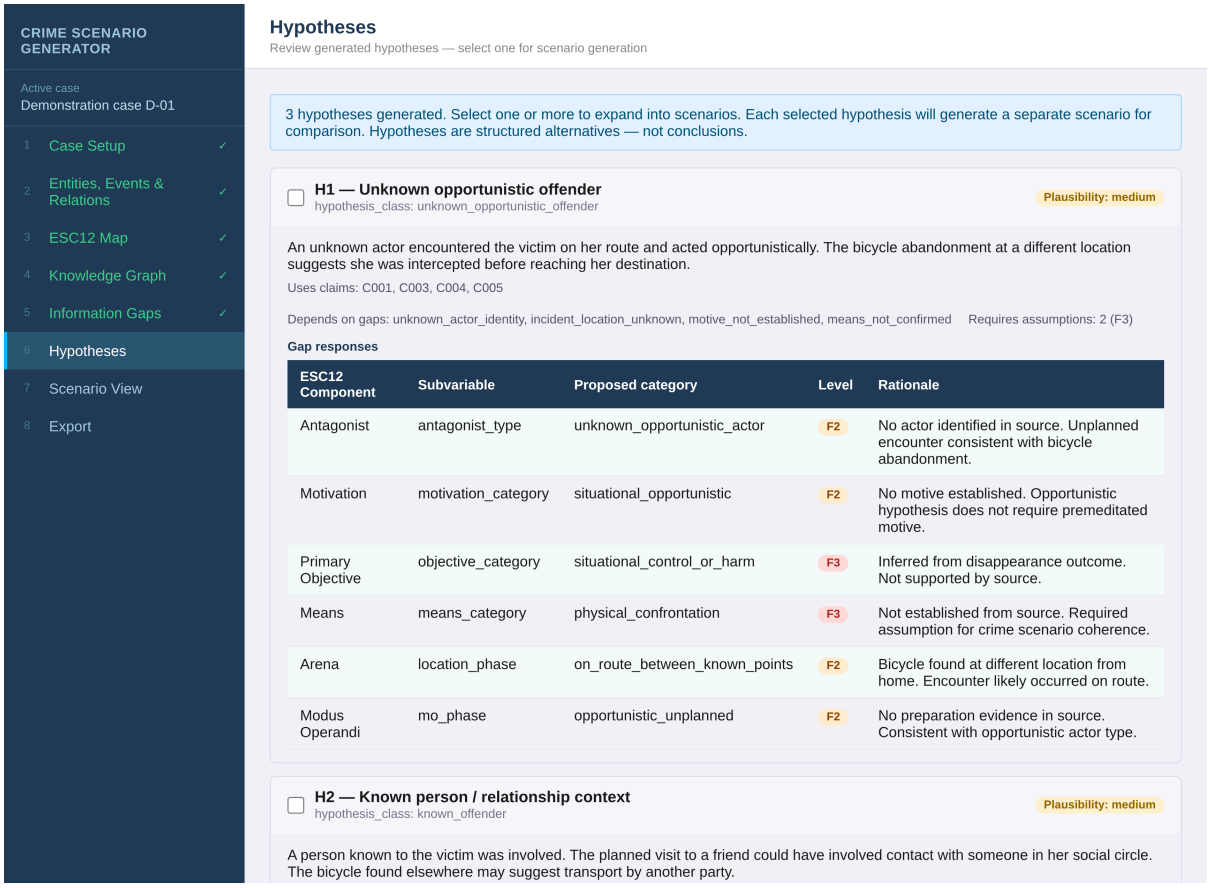


Figure 7. Generated alternative hypotheses in the application, shown with component-level evidential grading from established fact (F0) to assumption (F3) (Interface of the working prototype with demonstration data).

Digital Evidence Triage. The golden hours present a particular challenge in digital evidence management, as investigators must rapidly assess the relevance and reliability of data from mobile devices, computers, social media accounts, communication platforms, and internet-of-things devices. AI-based triage systems can prioritise evidence items based on relevance indicators, extract key information automatically, and flag potential connections between digital traces and physical evidence [12,32,38]. Anomaly detection techniques, such as those proposed by Studiawan and Sohel [59], can identify unusual patterns in forensic timelines that may indicate critical events or investigative leads.

Open-Source Intelligence (OSINT) Analysis. AI-powered OSINT tools can rapidly scan and analyse publicly available information, social media posts, news articles, online databases, and public records to generate leads and contextual intelligence during the golden hours [60]. These tools can identify connections between persons of interest, detect behavioural patterns, and provide background information that supports investigative decision-making. The integration of OSINT analysis with crime scene data and digital evidence creates a richer information landscape for golden-hour decision-making.

4.3. Evaluation of AI Techniques Against the Framework Criteria

Applying the five criteria defined in Section 3.5 to the four AI application domains examined above and the anchor definitions in Table 3 yields the author-assigned preliminary evidence-and-readiness assessment presented in Table 4. Ratings were assigned by the authors by consensus, applying those definitions to the evidence reviewed in Section 2 and to our previous experimental work [14,15,31], and the analysis of the case data. They

remain qualitative judgements rather than measurements, and are intended as a structured, transparent starting point for operational decision-making rather than as definitive scores. Two qualifications apply. First, the criteria are not operationally equivalent across domains: for 3D reconstruction, "accuracy" refers to geometric and visual fidelity under controlled conditions, not to the correctness of investigative conclusions; for evidence triage and OSINT it refers to retrieval precision and source reliability; and for narrative generation it refers to factual consistency and source-groundedness, which remain to be quantified. Second, "legal/governance readiness" is assessed for the intended use described in this paper under the legal framework discussed in Section 2.5, not as an intrinsic property of a technology; regulatory compliance as such is not an ordinal property of a technology and is therefore not what the rating expresses. The table should therefore be read as an evidence-and-readiness map rather than a validated comparative measurement, and obtaining ratings from independent experts is a planned next step.

Table 3. Anchor definitions used in the author-assigned preliminary evidence-and-readiness assessment (Table 4). Intermediate labels (Low–Medium, Medium–High) denote evidence spanning two adjacent anchors.

Criterion	High	Medium	Low
Speed	Actionable output within the golden hours using field-deployable capture	Output within roughly one working day	Requires days or laboratory processing
Accuracy	Quantitatively evaluated in published studies for the stated endpoint (for 3D reconstruction: geometric and visual fidelity under controlled conditions)	Plausible but probabilistic or interpretive output with limited validation	No systematic validation available
Explainability	Output directly verifiable by inspection (e.g., against the physical scene)	Output traceable to identifiable sources or rules	Opaque model reasoning
Legal/governance readiness	Low regulatory and ethical implementation burden for the defined use case under the applicable framework	Manageable privacy and bias risks requiring safeguards for the defined use case	High-risk processing without established safeguards
Adoption potential	Already embedded in operational police practice	Pilot-ready but requiring training and workflow integration	Experimental; judicial acceptability untested

Table 4. Author-assigned preliminary evidence-and-readiness assessment of AI application domains against the five golden-hour integration criteria. The assessors are the authors; where no direct validation evidence exists, the rating is given as "insufficient evidence".

AI Application Domain	Speed	Accuracy	Explainability	Legal/Governance Readiness	Adoption Potential	Evidence Basis
3D crime scene reconstruction (3DGS/NeRF)	High	High (controlled-condition fidelity)	High	High (documentation use)	Medium	[14,15,20]
Automated narrative generation	High	Insufficient evidence	Low–Medium	Medium	Low	Design sources [16,17]; no direct validation
Digital evidence triage	High	Medium–High	Medium	Medium	Medium	[12,38,59]
OSINT analysis	High	Medium	Medium	Medium	Medium–High	[60]

Rapid 3D reconstruction is the most mature of the four domains: capture and processing fit within golden-hour time constraints, reconstruction quality has been validated

experimentally in terms of image fidelity [14,15], which does not by itself establish forensic geometric measurement accuracy, and the outputs are visually reviewable against reference scene information, subject to reconstruction artefacts and validation. Its adoption potential is rated medium because equipment, training, and workflow integration still require organisational investment. Automated narrative generation offers high generation speed, although rapid text generation alone does not establish rapid delivery of usable investigative support, since the complete workflow, including source checking and the two-person review of provenance grades, has not been benchmarked; it directly targets tunnel vision, but its accuracy depends on probabilistic inference, its outputs are the least explainable of the four domains, and its judicial acceptability remains untested; it is therefore rated low on adoption potential at present. As a research prototype, it could be piloted as an analyst aid for guiding tasks during an investigation; any use of its outputs in an evidential role will take considerably longer and is to be tested in follow-up projects. Digital evidence triage achieves high speed and good accuracy, as reported in the cited studies and not benchmarked under golden-hour conditions in this study, for well-defined tasks such as prioritisation and anomaly detection [12,38,59], with moderate explainability and privacy-related compliance requirements. OSINT analysis is technically mature and already embedded in many police organisations [60], giving it the highest adoption-potential rating of the four domains in Table 4, ahead of 3D reconstruction, whose adoption still requires equipment, training, and workflow investment, but the reliability of open sources and data-protection constraints keep its accuracy and readiness ratings at medium. The anchor sets used in Tables 3 and 4 are heterogeneous and should be read accordingly: the accuracy anchors describe the amount and type of validation available rather than measured error magnitudes; the explainability anchors combine inspectability of outputs, traceability to sources, and understanding of model reasoning; and each rating is to be interpreted jointly with its evidence-basis entry, which identifies the specific task, the supporting evidence, and the operating conditions. Where the evidence basis is insufficient, this is recorded explicitly as insufficient evidence rather than expressed as an ordinal judgement.

4.4. Benefits for Evidence-Based Policing During the Golden Hours

The integration of AI technologies into golden-hour investigations offers several benefits aligned with evidence-based policing principles. Their evidential status differs and is indicated for each: some rest on demonstrated capabilities of the systems described above, while others are design goals of the framework whose operational effect has not yet been measured. First, as a design goal, AI is intended to reduce the dependency of investigative quality on individual expertise and experience by providing consistent, data-driven analytical support. This addresses a fundamental concern in current practice: the significant variability in investigative outcomes depending on the experience and cognitive style of the officer in charge [5–7].

Second, AI accelerates evidence processing and analysis: the reduction in capture time from hours to minutes, with processing completed within 1–2 h, is a demonstrated technical capability [14,15], whereas the automated extraction of digital evidence and the rapid generation of investigative leads through OSINT analysis are design goals that have not been benchmarked in the present study.

Third, as a design goal, AI supports the maintenance of evidential integrity during the golden hours. Digitally recorded 3D reconstructions provide an auditable, timestamped spatial record; a full chain of custody additionally requires procedural controls, integrity verification, metadata, audit logs, and documented handling. Automated logging of analytical processes supports transparency and auditability, both of which are essential requirements for judicial proceedings [12].

Fourth, as a demonstrated technical capability whose operational benefit remains to be evaluated, AI enhances collaborative investigation by enabling remote access to crime scene data, facilitating multi-agency coordination, and supporting real-time information sharing. This is particularly relevant in cases involving cross-border criminal networks, where rapid coordination between agencies is essential.

Fifth, as a design goal, AI can help distinguish between verified and unverified information during the golden hours, supporting investigators in prioritising reliable data and identifying information that requires further validation. By applying pattern recognition and consistency checking across multiple data sources, AI systems can flag potential discrepancies or contradictions that might otherwise go unnoticed in the pressure of the golden hours.

Sixth, as a design goal, AI can analyse many different data sources on site and in real time, whereas even a trained police officer cannot process such a volume and variety of data quickly and accurately. As the amount of digital data collected during criminal investigations continues to grow, the framework is designed to remain extensible to new data sources and analysis components.

4.5. Analysis of the Case Data

The case material was used at two levels. The corpus of 1589 open-source case records served as the source pool from which the coded sample was drawn; no corpus-level findings are claimed from it. The structural problems that motivate the framework, decision-making under incomplete information, evidence arriving out of chronological order, and the difficulty of maintaining an overview under organisational and external pressure, are documented in the investigative literature [3,4,7] and are illustrated concretely by the coded sample below. The qualitative review of the police archive material and the accompanying meetings with police personnel (Sections 3.3 and 3.4) pointed in the same direction and additionally suggested, as a contextual observation that cannot be verified from open sources, that there were instances in which digital-evidence processing times and specialist availability were perceived by practitioners as constraints during the golden hours; this observation is reported as context, not as a systematic finding.

The illustrative twenty-case sample introduced in Section 3.3 shows how these patterns manifest in individual investigations. The following counts describe this sample only and are not statistically generalisable. In three of the twenty cases, an initial investigative hypothesis or a dominant public narrative was later disproved: an assumed suicide that the autopsy revealed to be a homicide, a staged abduction account, and an incorrectly assumed honour-related motive. These two constructs are not equivalent and the coding treats them jointly only as instances in which an early explanation was later overturned: a publicly circulated motive that was subsequently rejected does not by itself demonstrate that investigators adopted that explanation or exhibited tunnel vision, and competing accounts adjudicated at trial do not necessarily identify a failure during the golden hours. In at least five cases, investigators or courts had to adjudicate explicitly between competing scenarios for the same evidence, for example a fall versus strangulation, or self-defence versus intentional violence. Digital traces, including camera footage, telecommunication data, GPS data, and online search history, played a substantial role in three cases. In three cases the decisive evidence emerged only months to years after the crime, including one breakthrough through large-scale DNA kinship screening thirteen years after the offence and in three cases linkage with other incidents was central to the investigation. Two cases in the sample remain unsolved. The full case-level coding, including the judicial outcome of each case, is provided in Supplementary Table S1, and the operational definitions of the coded categories in Supplementary Table S2; the notes to Table S1 state the aggregation

rules used for the plotted categories and distinguish a documented absence from a feature that the consulted sources simply do not report.

Figure 8 summarises the composition of the coded sample: the distribution of cases per decade, the recorded cause of death, the perpetrator–victim relation, and the judicial outcome. These descriptive statistics characterise the illustrative sample itself. They are not estimates for Dutch homicide cases in general.

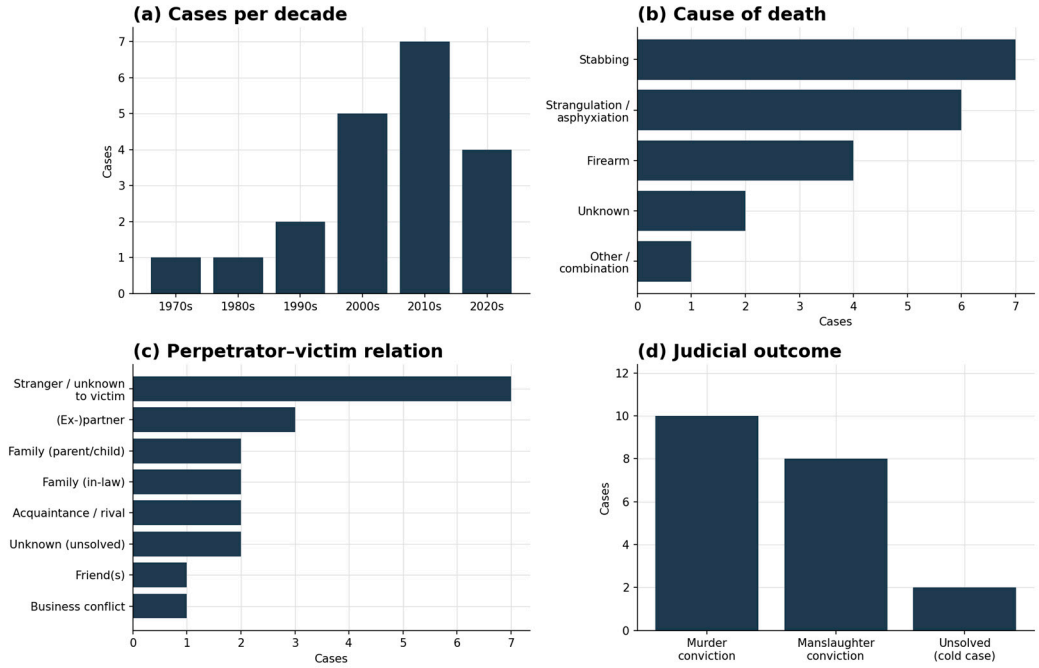


Figure 8. Descriptive statistics of the coded twenty-case sample: (**a**) cases per decade; (**b**) cause of death; (**c**) perpetrator–victim relation; (**d**) judicial outcome (n = 20; illustrative sample, not population estimates). Underlying case-level data and aggregation rules are provided in Supplementary Table S1 and the category definitions in Supplementary Table S2.

The Etten-Leur family homicide (case C-01 in Table 2) provides a worked example of where AI-supported evidence integration and scenario generation could have contributed during the golden hours. The surviving father initially claimed to have been abducted, while rumours of an honour-related motive circulated publicly; his account was disproved by confronting it with camera footage and telephone data. A system such as the one described in Section 4.2 is designed to support exactly this process: it would represent the abduction claim as one hypothesis among alternatives, including a domestic perpetrator, link every scenario component to the available digital traces, and grade each component by evidential support, so that inconsistencies between the claimed abduction and the camera and telephony timeline surface early and systematically rather than depending on ad hoc manual comparison, subject to the caveat that the early-phase availability of the digital inputs in this case is not documented (Table 5), so the timing of such a comparison in the actual investigation cannot be established retrospectively. In another case in the sample, indications of an acute threat, including a photograph of the suspect armed with two firearms, were available to the authorities before the crime but were not integrated into an effective intervention, underlining the operational cost of fragmented information flows [3,4]; this example concerns pre-offence threat assessment rather than post-incident golden-hour support and is included to illustrate evidence integration across fragmented reports, as distinguished in Table 5.

Table 5. Illustrative retrospective mapping of documented investigative challenges in five sample cases to candidate framework functions. Contributions are hypothesised and counterfactual.

Case	Documented Investigative Issue	Candidate Framework Function	Function Status	Early-Phase Availability of the Relevant Information	Hypothesised Contribution (Counterfactual)
C-01 (Etten-Leur, 2020)	Perpetrator's staged abduction account; public honour-motive rumours	Automated scenario generation; evidence integration	Prototype (text-based generation)	Partially documented (camera and telephony data obtained during the investigation; exact timing not documented)	Abduction claim held as one hypothesis among alternatives and confronted with the camera and telephony data as they arrived
C-03 (Dordrecht, 2011, cold case)	Initial suicide assumption; homicide indications emerged only at autopsy	Alternative-hypothesis generation at the scene phase	Prototype (text-based generation)	Partially documented (scene information available; autopsy findings followed later)	Homicide scenario retained alongside the suicide assumption at the scene phase
C-04 (Rijswijk, 2018)	Fall versus strangulation contested; three pathologists heard	Rapid 3D scene documentation	Demonstrated in prior work under controlled conditions	Documented (the scene itself was available for documentation)	Documented, auditable, revisitable spatial record of the scene to support adjudication between competing scenarios
C-05 (Oostburg, 2020)	Months-long search; GPS and shop traces ultimately decisive	Digital evidence triage	Design concept; not yet implemented	Not documented as available in the early phase	Systematic prioritisation of location and trace data within the triage workflow
C-06 (Rotterdam, 2018)	Threat indicators, including a photograph of the armed suspect, fragmented across separate reports	Evidence integration via linked case knowledge graph	Prototype (demonstration data)	Documented (the indicators pre-dated the offence)	Threat indicators linked and flagged as an integrated risk picture rather than isolated reports

These observations cannot be generalised beyond the material analysed, but they converge with the literature reviewed in Section 2: early explanations in real investigations were in some documented instances later overturned, which is consistent with, but does not by itself demonstrate, premature hypothesis fixation, since fixation would require case-specific evidence of sustained investigative commitment to the disproved explanation; digital traces played a decisive role in several sample cases; and critical evidence often arrived late and out of order. They therefore support the potential value of the framework proposed in Section 4.1, in which automated digital evidence triage addresses the processing bottleneck observed in the police data, and the systematic generation of alternative scenarios provides a structured counterweight to individual cognitive biases in the early phase of investigations. It bears repeating that the model generates research lines, not evidence: investigators use the generated narratives to search actively for supporting and contradicting information, newly collected data are entered back into the system, and ideally the set of narratives is then re-ranked, with the most plausible narrative at the top, together with the data that support this ranking and the information that ranks other narratives lower.

4.6. Retrospective Application of the Framework to Sample Cases

To move beyond the single worked example, the framework was mapped retrospectively onto five cases from the coded sample, selected so that each highlights a different framework component. For each case, Table 5 lists the investigative issue documented in the sources, the candidate framework function that addresses it, its demonstration status, whether the information relevant to the hypothesised contribution is documented as having been available in the early phase, and the contribution that function is hypothesised to have made; where early-phase availability is not documented, the contribution is phrased neutrally rather than as an earlier action. These assessments are counterfactual: the actual course of a supported investigation cannot be verified retrospectively. Their value is illustrative: they show that every major component of the framework corresponds to a concrete, documented problem in real Dutch investigations rather than to a hypothetical need; they do not establish that the framework would have changed the course or outcome of any investigation. The generated narratives are also organised around the central questions of every criminal investigation, the golden W questions (who, what, where, when, why, and how): each scenario component addresses one or more of these questions, and the documented issues in Table 5 illustrate how evidence and narrative are integrated around them.

5. Discussion

5.1. Implications for Evidence-Based Policing

This study suggests that AI technologies could meaningfully support evidence-based policing during the golden hours by addressing persistent challenges in investigative practice. The conceptual framework positions AI as an enabling layer that bridges digital documentation and higher-level interpretive tools, aligning with the broader trajectory of data-driven policing [11,25]. Embedded in golden-hour workflows, it supports Sherman's "triple-T" strategy at the operational level: targeting through evidence prioritisation, testing through scenario evaluation, and tracking through systematic documentation and audit trails. This mapping adapts, rather than directly operationalises, Sherman's framework: EBP concerns the use and testing of research evidence about policing practices at programme level, whereas the present framework concerns the handling of case evidence within a single investigation. Throughout this paper, "evidence-based" therefore refers to the research-to-practice paradigm, "data-driven" to the use of structured case data in decision support, and "AI-supported investigation" to the case-level mechanism proposed here; a data-driven tool is not automatically evidence-based, since it may rest on unvalidated models or biased historical data.

The integration of AI into the narrative approach advocated by Epskamp-Dudink [3] is particularly promising: AI can augment scenario reconstruction by systematically generating and evaluating multiple scenarios from the available evidence, enriching rather than replacing human judgment and reducing the risk of premature hypothesis fixation.

5.2. Comparison: Traditional vs. AI-Supported Golden-Hour Practices

Traditional golden-hour practices depend on the experience and cognitive capacity of the officer in charge. They benefit from human judgment, contextual understanding, and interpersonal skills, but are vulnerable to cognitive bias, information overload, and the limits of individual processing capacity under time pressure [7–9]. The model proposed here is designed to be more structured, documented, and traceable, because scenarios are generated from data and every scenario component carries its evidential support; at the same time, it may introduce its own sources of error and bias, originating from training data, source selection, model design, and human interpretation.

It can also process larger data volumes more rapidly, maintain systematic records of evidence and analysis, evaluate multiple hypotheses simultaneously, and provide more consistent analytical support across expertise levels, although output quality depends on data quality, model design, and use. However, AI systems currently lack the contextual understanding, ethical sensitivity, and interpersonal skills that characterise effective human investigators. The optimal approach therefore involves human-AI collaboration, where AI provides analytical support and scenario generation while human investigators retain control over interpretive judgments and investigative decisions [55]. The proposed model thus combines technology with human investigative work: if investigators do not engage seriously with a generated narrative and actively search for its pros and cons, its effectiveness will be low, making training an essential part of implementation.

5.3. Limitations and Challenges

Several significant limitations and challenges must be acknowledged. First, the reliability of AI-generated outputs remains a concern, particularly for scenario generation systems that involve interpretive inference rather than straightforward measurement or classification. AI systems can produce plausible but incorrect scenarios, and the persuasive quality of coherent narratives may lead investigators to accept AI-generated hypotheses without adequate critical evaluation [54].

These reliability concerns translate into concrete failure modes that any deployment must monitor: hallucinated entities or relationships, omitted or under-weighted evidence, corrupted or conflated timelines, duplicate identities across sources, unreliable or manipulated OSINT sources, adversarial manipulation of input data, model drift after updates, and incomplete input data during the earliest hours. Practical stopping rules follow from these failure modes: an AI-generated recommendation must be rejected, suspended, or escalated to manual review when it contradicts verified physical evidence, when its key components rest on assumptions (F3) rather than sourced claims, when the provenance of a component cannot be displayed, or when the input data are known to be materially incomplete. Because incomplete information and assumptions are inherent to early hypothesis generation, these conditions are not treated equivalently: contradiction of verified physical evidence, or provenance that cannot be displayed, requires rejection or abstention; materially incomplete inputs require suspension pending additional collection or review; and a hypothesis resting on flagged F3 assumptions may be cautiously retained as an explicitly speculative line of inquiry, provided the assumptions remain visible and the hypothesis is excluded from any evidential use.

Second, the explainability of AI systems is essential for judicial acceptance and investigative transparency, yet many advanced AI techniques, particularly deep learning models, operate as "black boxes" that cannot easily explain their reasoning processes [18,19]. Without explainable outputs, AI-generated analyses may face challenges in courtroom proceedings and may undermine rather than enhance investigative legitimacy.

Third, data quality and bias remain persistent risks: AI systems trained on historical crime data may perpetuate biases in policing patterns, evidence collection, and investigative priorities [39,40,42,44]. In golden-hour investigations, where rapid decisions on limited data are unavoidable, such biases could produce systematic errors in evidence prioritisation and scenario generation.

Fourth, organisational and cultural barriers to AI adoption should not be underestimated: besides technical infrastructure, integration requires training, change management, and cultural acceptance among investigators who may be sceptical of technology-driven approaches [27,28].

Fifth, the ethical and legal framework governing AI use in criminal investigations is still evolving. The EU AI Act classifies certain law-enforcement applications as "high-risk," imposing stringent requirements for transparency, human oversight, and accountability [57,58]; the applicable classification depends on the intended use of each component, for example scene documentation versus lead generation versus evidential assessment, and the processing of personal data in this context is governed by the Law Enforcement Directive rather than the GDPR [56]. Compliance with these requirements while maintaining the speed and flexibility needed for golden-hour operations presents a significant design challenge.

5.4. Future Directions

Several directions for future research emerge from this study. First, the framework has so far been explored retrospectively: the source pool and the coded twenty-case sample were used to identify documented investigative problems, and the framework was mapped onto five of these cases as an illustrative, counterfactual exercise (Sections 4.5 and 4.6); this does not constitute a validation of investigative effectiveness. The natural next step is prospective evaluation through pilot deployments in live investigative settings, involving experienced investigators and measuring impacts on decision quality, efficiency, and evidential outcomes. Any such evaluation will require an explicit operational cut-off for the golden hours, defined per case type (for example, the first eight or twenty-four hours), a time-stamped input packet containing only information demonstrably available at that cut-off, and outcome blinding, so that hindsight bias and leakage of case knowledge, including through the training data of foundation models, are controlled. Within that protocol, a retrospective comparison on solved cases, testing whether the model generates different and more plausible narratives than those actually pursued early in the investigation, offers a first quantifiable test of the generation process, since only solved cases provide a known outcome. A concise evaluation plan therefore comprises: a baseline, consisting of the hypotheses actually recorded early in each investigation or produced by investigators working without the tool; the time-restricted inputs described above; independent, blinded assessment of outputs by assessors not involved in system development; and measurable outcomes, including the rate of unsupported assertions in generated narratives, the omission of viable alternative scenarios, source-link accuracy, the human review workload per case, and the time to a defensible investigative decision. A solved case should not automatically be treated as complete ground truth for every component of its narrative: a conviction establishes some propositions to a judicial standard while leaving others disputed or unknown, so the protocol will specify in advance which propositions can be independently verified against documented evidence and will score disputed or unknown elements separately rather than as simply correct or incorrect.

Second, explainable AI techniques designed for investigative contexts require further attention: current approaches do not meet the requirements of criminal justice proceedings, where reasoning must be transparent, auditable, and comprehensible to non-technical audiences such as judges and juries.

Third, interdisciplinary training programmes for human-AI collaboration in golden-hour contexts should be developed and evaluated; the CSI effect, in which fictional portrayals inflate public expectations of forensic technology, underscores the need for training that conveys a realistic understanding of both the capabilities and the limitations of AI tools.

Finally, the framework is not limited to the Netherlands: homicide does not stop at national borders, and a European-oriented model, trained and evaluated on cases from multiple countries, is a direction we intend to pursue. Such a generalisation must account

for jurisdictional differences in criminal procedure, evidential admissibility, data access, police information systems, and national implementations of EU law.

6. Conclusions

This paper has explored the potential of artificial intelligence to enhance evidence-based and data-driven policing during the golden hours of criminal investigations. Through a structured review of the literature, analysis of real case data drawn from open sources and police records, and development of a conceptual framework, the study has identified several promising applications of AI for golden-hour investigations, including rapid 3D crime scene reconstruction, automated narrative generation, digital evidence triage, and open-source intelligence analysis. It bears repeating that this study is conceptual and exploratory: it does not validate investigative effectiveness, and all statements about operational benefit are hypotheses for the prospective evaluation outlined in Section 5.4.

The proposed conceptual framework integrates these AI capabilities within an evidence-based policing paradigm, positioning AI as a decision-support tool that augments rather than replaces human investigative judgment. The framework emphasises the principles of transparency, traceability, proportionality, and continuous evaluation, ensuring that AI functions as an investigation aid rather than a source of unquestionable truth.

The analysis suggests that AI could improve investigative efficiency and consistency and is hypothesised to mitigate the impact of cognitive bias and time pressure during the golden hours; none of these effects were measured in this study. At the same time, significant challenges remain regarding the reliability, explainability, and ethical compliance of AI systems in high-stakes investigative contexts. Addressing these challenges requires continued interdisciplinary collaboration between computer scientists, forensic scientists, legal scholars, ethicists, and law enforcement practitioners.

This work contributes to the theoretical understanding of AI-enhanced crime investigation by bridging the domains of evidence-based policing, scenario reconstruction, and AI-driven forensic analysis within the golden-hour context. The practical contribution lies in offering law enforcement agencies a structured approach to evaluating and integrating AI tools into early-stage investigations, supported by ethical guidelines and governance recommendations. As AI technologies continue to mature, the responsible integration of these tools into golden-hour practices holds promise for improving the quality and speed of criminal investigations, and potentially their fairness; demonstrating these effects requires prospective evaluation.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/forensicsci6040092/s1, Table S1: The coded twenty-case sample: case identifiers, descriptive variables, judicial outcomes, coded flags, and selection rationale; Table S2: Operational definitions of the coded flags used in Table S1 and in Section 4.5 of the main text.

Author Contributions: Conceptualization, D.R. and J.K.; methodology, D.R.; software, D.R.; investigation, D.R.; data curation, D.R.; writing (original draft preparation), D.R.; writing (review and editing), D.R. and J.K.; visualisation, D.R.; supervision, J.K.; funding acquisition, J.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received financial support from the research group Technologies for Criminal Investigations, part of Saxion University of Applied Sciences and the Police Academy of the Netherlands. This research is also part of the project "Regiodeal Stedendriehoek" funded by the Dutch National Government and "Hightech IGO" funded by Holland High Tech.

Institutional Review Board Statement: This study was conducted as part of a PhD project that received a positive ethical advice from the Computer & Information Sciences (CIS) Ethics Committee of the University of Twente (application no. 240519; positive advice issued on 26 November 2024).

The approval covers the involvement of human participants (interviews, focus groups, and workshops with investigative practitioners) and the use of pre-existing public homicide case data and police case information.

Informed Consent Statement: Informed consent was obtained from the practitioners who participated in the consultation activities, in accordance with the approved ethics application; no new data were collected from victims, suspects, or other private individuals, whose information was analysed only as documented in pre-existing public and police records.

Data Availability Statement: The underlying open-source case information used in this study is publicly accessible through Moordzaken.com, Moordatlas.nl, Uitspraken.nl, and Rechtspraak.nl, as described in Section 3.3. The compiled and coded dataset created by the authors from these publicly available sources, including the data underlying the analyses presented in this article, is available from the corresponding author upon reasonable request. The coded twenty-case table and the operational category definitions are provided as Supplementary Tables S1 and S2; per-case source links are not reproduced in S1, to avoid unnecessary identification of the individuals involved, and are available from the corresponding author; a confidential source register linking the case identifiers to the underlying public documentation has been provided to the editorial office for verification. The complementary dataset provided by the Police of the Netherlands is not publicly available due to confidentiality, data-governance, and access restrictions imposed by the data owner. The authors are not authorised to redistribute this dataset or provide access to third parties.

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

References

1. Delémont, O.; Lock, E.; Ribaux, O. Forensic Science and Criminal Investigation. In *Encyclopedia of Criminology and Criminal Justice*; Springer: New York, NY, USA, 2014; pp. 1754–1763. [CrossRef]
2. Bitzer, S.; Albertini, N.; Lock, E.; Ribaux, O.; Delémont, O. Utility of the clue—From assessing the investigative contribution of forensic science to supporting the decision to use traces. *Sci. Justice* **2015**, *55*, 509–513. [CrossRef] [PubMed]
3. Epskamp-Dudink, C. *Van Watskeburt tot Whodunit: Stap voor Stap Scenario's Reconstrueren*; Boom Criminologie: Den Haag, The Netherlands, 2023.
4. Epskamp-Dudink, C.; Winter, J.M. Benefits of scenario reconstruction in cold case investigations. *J. Crim. Psychol.* **2020**, *10*, 65–78. [CrossRef]
5. Swanson, C.; Taylor, R.W.; Territo, L.; Fox, B.; Chamelin, N. *Criminal Investigation*; McGraw-Hill: New York, NY, USA, 2023.
6. Buckles, T. *Crime Scene Investigation, Criminalistics and the Law*; Cengage Learning: New York, NY, USA, 2007.
7. Groenendaal, J.; Helsloot, I. Tunnel vision on tunnel vision? A preliminary examination of the tension between precaution and efficacy in major criminal investigations in the Netherlands. *Police Pract. Res.* **2015**, *16*, 224–238. [CrossRef]
8. Findley, K.A. Tunnel vision. In *Conviction of the Innocent: Lessons from Psychological Research*; American Psychological Association: Washington, DC, USA, 2012; pp. 303–323. [CrossRef]
9. Elaad, E. Tunnel Vision and Confirmation Bias Among Police Investigators and Laypeople in Hypothetical Criminal Contexts. *Sage Open* **2022**, *12*, 21582440221095022. [CrossRef]
10. Salet, R.; Terpstra, J. Critical Review in Criminal Investigation: Evaluation of a Measure to Prevent Tunnel Vision. *Policing* **2014**, *8*, 43–50. [CrossRef]
11. Sherman, L.W. The Rise of Evidence-Based Policing: Targeting, Testing, and Tracking. *Crime Justice* **2013**, *42*, 377–451. [CrossRef]
12. Dunsin, D.; Ghanem, M.C.; Ouazzane, K.; Vassilev, V. A comprehensive analysis of the role of artificial intelligence and machine learning in modern digital forensics and incident response. *Forensic Sci. Int. Digit. Investig.* **2024**, *48*, 301675. [CrossRef]
13. Carew, R.M.; French, J.; Morgan, R.M. 3D forensic science: A new field integrating 3D imaging and 3D printing in crime reconstruction. *Forensic Sci. Int. Synerg.* **2021**, *3*, 100205. [CrossRef] [PubMed]
14. Rangelov, D.; Waanders, S.; Waanders, K.; van Keulen, M.; Miltchev, R. Impact of Camera Settings on 3D Reconstruction Quality: Insights from NeRF and Gaussian Splatting. *Sensors* **2024**, *24*, 7594. [CrossRef] [PubMed]
15. Rangelov, D.; Waanders, S.; Waanders, K.; van Keulen, M.; Miltchev, R. Impact of Data Capture Methods on 3D Reconstruction with Gaussian Splatting. *J. Imaging* **2025**, *11*, 65. [CrossRef] [PubMed]
16. Ugai, T.; Koyanagi, Y.; Nishino, J.F. A Logical Approach to Criminal Case Investigation. *arXiv* **2024**, arXiv:2402.08284.
17. De Kock, P.A.M.G. Anticipating Criminal Behaviour: Using the Narrative in Crime-Related Data. Ph.D. Thesis, Tilburg University, Tilburg, The Netherlands, 2014.

18. Bharati, R.K. Ethical Implications of AI in Criminal Justice: Balancing Efficiency and Due Process. *Res. Rev. Int. J. Multidiscip.* **2024**, *9*, 93–105. [CrossRef]
19. John, A.M.; Anitha, M.U.; Panachakel, J.T. Ethical Challenges of Using Artificial Intelligence in Judiciary. *arXiv* **2025**, arXiv:2504.19284.
20. Rangelov, D.; Knotter, J.; Miltchev, R. 3D Reconstruction in Crime Scenes Investigation: Impacts, Benefits, and Limitations. *Lect. Notes Netw. Syst.* **2024**, *1065*, 46–64. [CrossRef]
21. Noiret, S.; Lumetzberger, J.; Kampel, M. Bias and Fairness in Computer Vision Applications of the Criminal Justice System. In Proceedings of the 2021 IEEE Symposium Series on Computational Intelligence, Orlando, FL, USA, 5–7 December 2021. [CrossRef]
22. Fletcher, C. *Every Contact Leaves a Trace: Crime Scene Experts Talk About Their Work from Discovery Through Verdict*; St. Martin's Press: New York, NY, USA, 2006.
23. Locard, E. *L'Enquête Criminelle et les Méthodes Scientifiques*; Flammarion: Paris, France, 1920.
24. Barone, P.M.; Di Luise, E. A Multidisciplinary Approach to Crime Scene Investigation: A Cold Case Study and Proposal for Standardized Procedures in Buried Cadaver Searches over Large Areas. *Forensic Sci.* **2025**, *5*, 34. [CrossRef]
25. Sherman, L.W.; Murray, A. Evidence-Based Policing: From Academics to Professionals. *Int. Crim. Justice Rev.* **2015**, *25*, 7–10. [CrossRef]
26. Knutsson, J.; Tompson, L. *Advances in Evidence-Based Policing*; Routledge: London, UK, 2017.
27. Telep, C.W.; Lum, C. The Receptivity of Officers to Empirical Research and Evidence-Based Policing: An Examination of Survey Data From Three Agencies. *Police Q.* **2014**, *17*, 359–385. [CrossRef]
28. van Dijk, A.; Hoogewoning, F.; Punch, M. *What Matters in Policing? Change, Values and Leadership in Turbulent Times*; Policy Press: Bristol, UK, 2015.
29. Greene, J.R. New Directions in Policing: Balancing Prediction and Meaning in Police Research. *Justice Q.* **2014**, *31*, 193–228. [CrossRef]
30. Presser, L.; Sandberg, S. *Narrative Criminology*; NYU Press: New York, NY, USA, 2015.
31. Rangelov, D.; Knotter, J. Between Insight and Influence: Ethical Constraints of AI-Generated Crime Scene Scenarios. In *Proceedings of the International Conference on Intelligent Systems*; Lecture Notes in Networks and Systems; Springer: Cham, Switzerland, 2025.
32. Hu, K.; Li, L.; Tao, X.; Velásquez, J.D.; Delaney, P. Information fusion in crime event analysis: A decade survey on data, features and models. *Inf. Fusion* **2023**, *100*, 101904. [CrossRef]
33. Farber, S. AI as a decision support tool in forensic image analysis: A pilot study on integrating large language models into crime scene investigation workflows. *J. Forensic Sci.* **2025**, *70*, 932–943. [CrossRef] [PubMed]
34. Sessa, F.; Chisari, M.; Esposito, M.; Guardo, E.; Di Mauro, L.; Salerno, M.; Pomara, C. Advancing Diagnostic Tools in Forensic Science: The Role of Artificial Intelligence in Gunshot Wound Investigation—A Systematic Review. *Forensic Sci.* **2025**, *5*, 30. [CrossRef]
35. Galanakis, G.; Zabulis, X.; Evdaimon, T.; Fikenscher, S.-E.; Allertseder, S.; Tsikrika, T.; Vrochidis, S. A Study of 3D Digitisation Modalities for Crime Scene Investigation. *Forensic Sci.* **2021**, *1*, 56–85. [CrossRef]
36. Nigliaccio, S.; Fontana, D.; Di Vita, E.; Piraino, M.; Messina, P.; Argo, A.; Zerbo, S.; Albano, D.; Cumbo, E.; Scardina, G.A. Post-Mortem Animal Bite Mark Analysis Reimagined: A Pilot Study Evaluating the Use of an Intraoral Scanner and Photogrammetry for Forensic 3D Documentation. *Forensic Sci.* **2025**, *5*, 39. [CrossRef]
37. Hajare, O.; Thalor, M. AI-Based Crime Scene Simulation through 3D Image Processing and Semantic Segmentation. *Int. J. Sci. Res. Eng. Trends* **2025**, *11*.
38. Bhardwaj, A.; Choudhary, S.K. AI Based Decision Support System for Cyber Forensics Investigations. In Proceedings of the 2024 IEEE 4th International Conference on ICT in Business Industry and Government, Indore, India, 13–14 December 2024. [CrossRef]
39. Ziosi, M.; Pruss, D. Evidence of What, for Whom? The Socially Contested Role of Algorithmic Bias in a Predictive Policing Tool. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*; Association for Computing Machinery: Rio de Janeiro, Brazil, 2024. [CrossRef]
40. O'Donnell, R.M. Challenging Racist Predictive Policing Algorithms Under the Equal Protection Clause. *N. Y. Univ. Law Rev.* **2019**, *94*, 544.
41. Lum, K.; Isaac, W. To predict and serve? *Significance* **2016**, *13*, 14–19. [CrossRef]
42. Richardson, R.; Schultz, J.M.; Crawford, K. Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice. *N. Y. Univ. Law Rev. Online* **2019**, *94*, 15–55.
43. Brayne, S. Big Data Surveillance: The Case of Policing. *Am. Sociol. Rev.* **2017**, *82*, 977–1008. [CrossRef] [PubMed]
44. Meijer, A.; Wessels, M. Predictive Policing: Review of Benefits and Drawbacks. *Int. J. Public Adm.* **2019**, *42*, 1031–1039. [CrossRef]
45. Ensign, D.; Friedler, S.A.; Neville, S.; Scheidegger, C.; Venkatasubramanian, S. Runaway Feedback Loops in Predictive Policing. In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*; PMLR: London, UK, 2018; Volume 81, pp. 160–171.
46. Ferguson, A.G. Policing Predictive Policing. *Wash. Univ. Law Rev.* **2017**, *94*, 1109–1189.

47. Bennett Moses, L.; Chan, J. Algorithmic prediction in policing: Assumptions, evaluation, and accountability. *Polic. Soc.* **2018**, *28*, 806–822. [CrossRef]
48. Dressel, J.; Farid, H. The accuracy, fairness, and limits of predicting recidivism. *Sci. Adv.* **2018**, *4*, eaao5580. [CrossRef] [PubMed]
49. Skitka, L.J.; Mosier, K.L.; Burdick, M. Does automation bias decision-making? *Int. J. Hum. Comput. Stud.* **1999**, *51*, 991–1006. [CrossRef]
50. Goddard, K.; Roudsari, A.; Wyatt, J.C. Automation bias: A systematic review of frequency, effect mediators, and mitigators. *J. Am. Med. Inform. Assoc.* **2012**, *19*, 121–127. [CrossRef] [PubMed]
51. Parasuraman, R.; Riley, V. Humans and automation: Use, misuse, disuse, abuse. *Hum. Factors* **1997**, *39*, 230–253. [CrossRef]
52. Parasuraman, R.; Manzey, D.H. Complacency and bias in human use of automation: An attentional integration. *Hum. Factors* **2010**, *52*, 381–410. [CrossRef] [PubMed]
53. Cummings, M.L. Automation bias in intelligent time critical decision support systems. In Proceedings of the AIAA 1st Intelligent Systems Technical Conference, Chicago, IL, USA, 20–22 September 2004. [CrossRef]
54. Alon-Barkat, S.; Busuioc, M. Human–AI Interactions in Public Sector Decision Making: 'Automation Bias' and 'Selective Adherence' to Algorithmic Advice. *J. Public Adm. Res. Theory* **2023**, *33*, 153–169. [CrossRef]
55. Park, P.S.; Goldstein, S.; O'Gara, A.; Chen, M.; Hendrycks, D. AI deception: A survey of examples, risks, and potential solutions. *Patterns* **2024**, *5*, 100988. [CrossRef] [PubMed]
56. European Parliament and Council of the European Union. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). In *Official Journal of the European Union*; Publications Office of the European Union: Luxembourg, 2016; pp. 89–131. Available online: http://data.europa.eu/eli/dir/2016/680/oj (accessed on 10 September 2026).
57. European Parliament and Council of the European Union. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). In *Official Journal of the European Union*; Publications Office of the European Union: Luxembourg, 2024. Available online: http://data.europa.eu/eli/reg/2024/1689/oj (accessed on 10 September 2026).
58. Novelli, C.; Casolari, F.; Hacker, P.; Spedicato, G.; Floridi, L. Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity. *Comput. Law Secur. Rev.* **2024**, *55*, 106066. [CrossRef]
59. Studiawan, H.; Sohel, F. Anomaly detection in a forensic timeline with deep autoencoders. *J. Inf. Secur. Appl.* **2021**, *63*, 103002. [CrossRef]
60. Trottier, D. Open source intelligence, social media and law enforcement: Visions, constraints and critiques. *Eur. J. Cult. Stud.* **2015**, *18*, 530–547. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.