Trustworthy AI and Mixed Reality in Police Interventions: Challenges and Opportunities

ANDREAS BRÄNNSTRÖM, Department of Computing Science, Umeå University, Sweden
EDUARDO GARCÍA LAREDO, International University of La Rioja (UNIR), Spain
BERNAT VIVOLAS JORDÀ, Institute for Public Security of Catalunya, Spain
LOLA VALLES, University of Barcelona, Spain
JONAS HANSSON, Unit of Police Work, Umeå University, Sweden
EMILI MARTINEZ CAÑAVERAS, Institute for Public Security of Catalunya, Spain
ANDERS SCHOGSTER, Unit of Police Work, Umeå University, Sweden
DAVID MARTIN-MONCUNILL, Comet Global Innovation, Spain and Camilo José Cela University (UCJC), Spain
JUAN CARLOS NIEVES*, Department of Computing Science, Umeå University, Sweden

Designing Artificial Intelligence (AI)-based interactive systems for law enforcement poses unique socio-technical and ethical challenges, particularly when such systems must support real-time decision-making in dynamic, high-stakes environments. Despite their potential, AI-supported interactive systems in policing require carefully elicited domain-specific requirements to ensure effective use while being Ethical by Design. However, methods for such requirement elicitation remain limited. This paper presents a participatory approach for identifying the requirements of AI-driven Mixed Reality (MR) systems in law enforcement contexts. The introduced methodology builds on two of our previous EU projects: the Erasmus+ project "Trustworthy AI", which provided educational material to teach key principles of Trustworthy AI, and AI4EU, which developed an abbreviated assessment tool for evaluating AI systems. In collaboration with police education units and law enforcement agencies in Sweden and Catalonia, we conducted a multi-phase study involving two preparatory workshops—one focused on educating participants in Trustworthy AI, and another involving hands-on MR use in standard police training scenarios. This reflects the view that it is not enough to simply ask people about new technologies—they must also be educated to critically assess their implications. After the workshops, we collected structured feedback using quantitative and qualitative methods. To analyze risk levels of the elicited requirements, we applied the AI4EU-based assessment tool. Our findings highlight key challenges and opportunities for designing AI-based systems with MR interfaces that enhance decision-making in real-time police operations while ensuring transparency, safety, and human oversight.

*Corresponding author.

Authors' Contact Information: Andreas Brännström, ORCID: 0000-0001-9379-4281, andreas.brannstrom@umu.se, Department of Computing Science, Umeå University, Umeå, Sweden; Eduardo García Laredo, ORCID: 0000-0002-7250-0456, eduardo.garcialaredo@unir.net, International University of La Rioja (UNIR), Logroño, Spain; Bernat Vivolas Jordà, ORCID: 0009-0005-1340-1318, bvivolasj@gencat.cat, Institute for Public Security of Catalunya, Spain; Lola Valles, ORCID: 0000-0001-5899-9927, lvalles@gencat.cat, University of Barcelona, Spain; Jonas Hansson, ORCID: 0000-0001-6113-414X, jonas.hansson@umu.se, Unit of Police Work, Umeå University, Umeå, Sweden; Emili Martinez Cañaveras, ORCID: 0009-0001-6219-9575, emili.martinez@mossos.cat, Institute for Public Security of Catalunya, Spain; Anders Schogster, ORCID: 0009-0004-0891-9254, anders.schogster@umu.se, Unit of Police Work, Umeå University, Umeå, Sweden; David Martin-Moncunill, ORCID: 0000-0003-2422-9005, d.martin@es.comet.technology, Comet Global Innovation, Barcelona, Catalonia, Spain and Camilo José Cela University (UCJC), Villafranca del Castillo, Spain; Juan Carlos Nieves, ORCID: 0000-0003-4072-8795, juan.carlos.nieves@umu.se, Department of Computing Science, Umeå University, Umeå, Sweden.

DOI: 10.1613/jair.1.19461

JAIR Track: AI and Society Track

JAIR Associate Editor: Ulises Cortés

JAIR Reference Format:
Andreas Brännström, Eduardo García Laredo, Bernat Vivolas Jordà, Lola Valles, Jonas Hansson, Emili Martinez Cañaveras, Anders Schogster, David Martin-Moncunill, and Juan Carlos Nieves. 2026. Trustworthy AI and Mixed Reality in Police Interventions: Challenges and Opportunities. *Journal of Artificial Intelligence Research* 86, Article 12 (June 2026), 30 pages. DOI: 10.1613/jair.1.19461

1 Introduction

Artificial Intelligence (AI) has emerged as a powerful tool for augmenting human capabilities in complex domains. By enabling predictive analytics, automated reasoning, and multi-source data integration, AI systems have transformed sectors such as healthcare (Shaheen 2021), finance (Cao 2022), transportation (Abduljabbar et al. 2019), and security (Perez-Cerrolaza et al. 2024). At the same time, interactive technologies like Mixed Reality (MR) (Speicher et al. 2019) provide new modalities for delivering these capabilities to users in real time. MR headsets—such as Microsoft HoloLens—enable hands-free interfaces that layer digital content directly onto the user's visual field. This opens up opportunities to present information in more context-sensitive ways (Rokhsaritalemi et al. 2020).

Law enforcement is a safety-critical domain characterized by rapid decision-making, limited information, and dynamic public environments. Whether responding to routine traffic stops (Woods 2018) or high-risk situations such as active shooter incidents (Blair et al. 2019), police officers operate under time pressure in situations that may involve the safety of suspects, victims, bystanders, and fellow officers, while remaining accountable to legal standards and institutional protocols. In the context of law enforcement, AI has been explored as a tool for supporting policing and digital investigations, in applications such as predictive policing (Leese 2021), traffic control analysis (Aouedi et al. 2022), evidence analysis (Costantini et al. 2019), automated forensics in mobile devices (Anglano et al. 2021), and bias-neutralization in digital investigations (Renaud et al. 2021). AI-based systems offer the potential to enhance situational awareness for police officers by analyzing past events, recognizing emerging threats, and providing decision-support in high-risk interventions. When integrated into MR interfaces, such insights can be presented in a way that is immediately accessible, hands-free, and embedded within the operational context (Apostolakis et al. 2021; Vossers et al. 2024). The combination of AI and MR thus holds strong potential for improving officers' situation awareness, tactical coordination, and public safety during police interventions.

Yet, the design and deployment of such intelligent systems in law enforcement raise substantial socio-technical and ethical challenges. AI outputs may influence high-stakes decisions, requiring systems that are transparent, fair, and robust under uncertainty. Policing, in particular, demands sensitivity to ethical concerns such as privacy, accountability, and non-discrimination—especially in public-facing applications. Moreover, interventions unfold in fast-changing, multi-actor environments, where system reliability, user autonomy, and cognitive demands must all be carefully considered. To ensure that such systems align with both operational realities and public values, their development must be guided by domain-specific requirements that reflect the needs and constraints of the setting in which they are deployed.

A central premise of contemporary approaches to Trustworthy and Responsible Artificial Intelligence is that ethical considerations must be addressed proactively as part of the system design process, rather than retrospectively through regulation or post-deployment controls. This perspective is commonly referred to as Ethics by Design (Dignum, Baldoni, et al. 2018) and emphasizes the systematic integration of ethical, legal, and societal values into the technical development lifecycle of AI systems. Within software engineering, this places particular importance on the early phases of development, and especially on requirements elicitation, where system goals, constraints, data assumptions, and acceptable trade-offs are first articulated. Decisions made at this stage—such as which data sources are permissible, how system autonomy is bounded, what forms of human oversight are required, and which risks are deemed acceptable—fundamentally shape the ethical properties of the resulting system. For AI-based systems operating in safety-critical and socially sensitive domains, such as law enforcement, requirements elicitation therefore becomes a key mechanism for operationalizing ethical principles, translating high-level guidelines for Trustworthy AI into concrete, context-specific system requirements that can guide subsequent design, implementation, and evaluation. While requirement elicitation is recognized as a foundational phase for aligning system behavior with legal, ethical, and functional constraints, considering requirements of Trustworthy AI (Dignum, Nieves, et al. 2021; European Commission 2019), existing tools and methods are often generic, underdeveloped, or inconsistently applied in real-world projects (Guizzardi et al. 2020; Krafft et al. 2020). A recent survey shows that although ICT practitioners frequently use standard techniques such as interviews and stakeholder meetings, few report applying elicitation methods tailored to the unique challenges of AI systems—including those related to fairness, explainability, and data governance (Sousa Silva et al. 2022). Ethical requirements such as equity, inclusion, and accountability are also frequently under-addressed in early development (Bibal et al. 2021; Jobin et al. 2019).

The aforementioned challenges and opportunities, situated in the context of law enforcement, give rise to the following research questions:

- **Q1:** Should AI-based interactive systems be used during police interventions?
- **Q2:** What are the requirements for AI-based interactive systems to be effective and ethical in real-time policing?
- **Q3:** What AI ethics risks arise in the use of AI-based interactive systems in law enforcement?

In this work, we develop, apply, and analyze an overarching methodology for eliciting requirements for AI-based interactive systems in law enforcement. This methodology incorporates components from our two previous EU-funded projects: educational materials on Trustworthy AI from the Erasmus+ project "Trustworthy AI" [1], used to prepare participants before the study, and the "abbreviated assessment list" from the project "AI4EU" (Dignum, Nieves, et al. 2021), used to analyze the elicited requirements. Our study was conducted in collaboration with computing science researchers, Police Education Units, and law enforcement agencies in Sweden and Catalonia (Spain)[2], and resulted in both methodological insights and empirically grounded findings. Recognizing that participants bring diverse backgrounds and varying levels of familiarity with AI technologies and ethical concepts, the study begins with an introductory workshop on Trustworthy AI. This preparatory step helps ensure that participants interpret the study materials and questions in a consistent conceptual frame, enabling more coherent and analyzable feedback during the subsequent mixed reality user study. In the next phase, participating police officers engaged in a standard operational training exercise—while wearing mixed reality headsets—after which a combination of quantitative and qualitative methods were applied to analyze participants' experiences. A customized survey based on the well-known Technology Acceptance Model (TAM) (Marangunić and Granić 2015) extended with trust-based measures, called the Degree of Operational Confidence in Artificial Systems (POTDAI) (Martín-Moncunill et al. 2024), to collect structured feedback on perceived usefulness, ease of use, and system trustworthiness. We further conducted a so-called Pluralistic Walkthrough (Bias 1994)—a structured usability evaluation method that involves future users reviewing early design concepts—to elicit qualitative insights into usability, preferred information delivery, and system expectations. Together, these components form a systematic and participatory approach for identifying ethically grounded and operationally relevant requirements for AI-based tools in policing (see Figure 1).

[1]EU Trustworthy AI project: https://www.trustworthyaiproject.eu/
[2]Project website: https://www.humane-ai.eu/project/tmp-121/

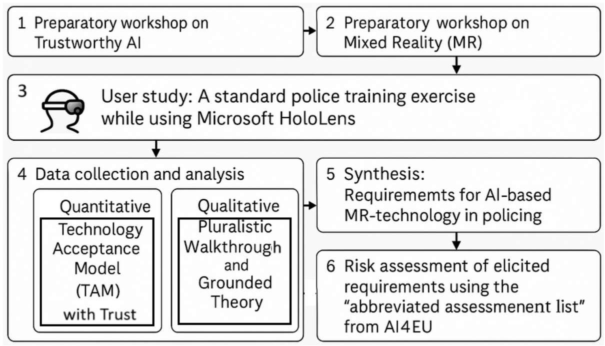


Fig. 1. Study Design and Methodology: Requirement Elicitation for AI-based MR-support in Policing.

Our findings contribute to the identification and organization of system requirements in law enforcement, helping guide the design of AI-based MR systems that enhance situational awareness while preserving transparency, trust, and officer autonomy.

The rest of this paper is structured as follows. Section 2 provides an overview of the study participants and outlines the overall study design. Section 3 describes the two preparatory workshops on Trustworthy AI and Mixed Reality. Section 4 details the user study conducted during a standard police training exercise. Section 5 explains our requirement elicitation methods, including both quantitative (TAM and POTDAI) and qualitative (Pluralistic Walkthrough) techniques. Section 6 presents the results of the quantitative analysis, and Section 7 reports the qualitative findings. Section 8 discusses these results in the context of related work and broader implications, and Section 9 concludes the paper.

2 Participants and Study Overview

Two separate studies were conducted as part of this research, one in Catalonia (Spain) and one in Sweden, both involving experienced police officers at national police training institutions. The goal was to explore how AI-supported MR technologies could be meaningfully integrated into real-world police operations, while eliciting both functional and ethical system requirements from experienced practitioners. Each study followed an identical structure consisting of the following main phases (see Figure 1):

(1) Preparatory workshop on Trustworthy AI
(2) Preparatory workshop on Mixed Reality (MR)
(3) User study: A standard police training exercise while using Microsoft HoloLens
(4) Data collection and analysis:
  (a) Quantitative: Technology Acceptance Model (TAM) (Marangunić and Granić 2015) with Trust (Martín-Moncunill et al. 2024)
  (b) Qualitative: Pluralistic Walkthrough (Bias 1994) and Grounded Theory (Oktay 2012)
  (c) Synthesis: Requirements for AI-based MR-technology in policing
  (d) Risk assessment of elicited requirements using the "abbreviated assessment list" from AI4EU (Dignum, Nieves, et al. 2021).

In Catalonia, the study was conducted at the Police School of Catalonia in October 2023. All instructors were invited to participate in a two-day workshop series organized by the research team. Twenty officers attended both days and gave informed consent to be included in the study. The group had an even gender distribution, with ten male and ten female participants, aged 31 to 50, and with 12 to 19 years of service.

In Sweden, the study was carried out at the Unit of Police Work at Umeå University in February 2024. Nineteen instructors volunteered to participate in the two-day study, which followed the same structure as the Catalan study. The Swedish group consisted of sixteen male and three female officers, aged 30 to 59, with service lengths ranging from 9 to 37 years.

Demographic details for both groups are summarized in Table 1.

Table 1. Participants: Catalan and Swedish Police Officers

	Catalan police officers (N=20)			**Swedish police officers (N=19)**		
	Male	**Female**	**Female%**	**Male**	**Female**	**Female%**
Sex	10	10	50%	16	3	18.75%
	Min	**Max**	**Mean**	**Min**	**Max**	**Mean**
Age	31	50	40.90	30	59	48.70
Years of Service	12	19	15.25	9	37	19.79

This methodology aimed to support informed and experience-based feedback from participants. After completing the police exercise, officers were asked to reflect on what kinds of information should be presented through MR, when it would be most useful, and how it should be delivered to support decision-making during interventions. Each phase of the methodology is described in the sections that follow.

3 Preparatory Workshops

To ensure a common foundation among participants prior to the user study, we conducted two preparatory workshops. These workshops aimed to build both ethical and technological literacy relevant to the study. The first focused on introducing participants to the key principles of Trustworthy AI, while the second offered a hands-on experience with Mixed Reality (MR) technology using Microsoft HoloLens. All participants were active law enforcement officers, and the workshops were conducted in collaboration with police education units in Sweden and Catalonia.

3.1 Trustworthy AI Workshop

Prior to the user study, participants attended a preparatory workshop introducing the concept of Trustworthy Artificial Intelligence (AI). The aim was to create a shared reference point for reflecting on ethical issues in AI development and deployment, particularly in the context of law enforcement. Rather than providing instruction in applied ethics, the workshop was designed as a neutral and participatory space for open-ended group exploration.

The session was structured using materials from the *Trustworthy AI project*[3], supported by the European Commission. Developed to support AI ethics education across higher education, the project provides general-purpose resources—including explanatory videos and a card-based group activity—which we adapted for use in a policing context. Figure 2 shows participants being presented videos on Trustworthy AI.

Each of the seven principles defined in the European Commission's Ethics Guidelines for Trustworthy AI (European Commission 2019)—namely: human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity and fairness, societal and environmental well-being, and accountability—was introduced through short videos developed by the *Trustworthy AI project*.

Following the video introductions, participants engaged in the card game activity in small groups. The cards presented discussion prompts involving stakeholders, technologies, risks, and scenarios related to AI deployment, centered on a case in the domain of law enforcement. The activity aimed to spark collaborative reasoning about

[3]EU Trustworthy AI project: https://www.trustworthyaiproject.eu/

Fig. 2. Trustworthy AI workshop - Videos on AI Ethics Guidelines

how ethical principles might be interpreted in real-world law enforcement contexts, based on participants' own professional experience. Care was taken to avoid steering participants toward particular ethical positions or design assumptions. Instead, the focus was on providing space for unstructured dialogue and grounded reflection. Positioned before the MR-based user study, the workshop aimed to support more thoughtful engagement with the role of AI technologies in police operations.

3.2 Mixed Reality Workshop

The second preparatory workshop introduced participants to Mixed Reality (MR) technologies through direct engagement with Microsoft HoloLens devices (see Figure 3). This session aimed to provide participants with a tangible sense of how digital content can be spatially anchored and interactively overlaid onto physical environments. Rather than demonstrating a specific prototype, the focus was on enabling experiential understanding of MR capabilities, supporting participants in imagining future applications of such technologies within their own operational context.

Fig. 3. Mixed-Reality Workshop to explore MR-functionalities in Microsoft HoloLens

The workshop consisted of two main parts. In the first part, participants explored different modalities for receiving information through MR. Using Microsoft Remote Assist[4] installed on Microsoft HoloLens, researchers were able to observe the participants' live point of view on an external screen. This setup allowed facilitators to remotely place instructional elements—such as directional arrows on the ground or walls, or holograms

[4]https://learn.microsoft.com/en-us/dynamics365/mixed-reality/remote-assist/

representing people or objects—within the participants' field of vision. Participants were then instructed to follow these MR cues and navigate the environment accordingly. This part of the session demonstrated how MR can be used to deliver context-sensitive instructions or augment situational awareness in real time.

In the second part of the workshop, participants interacted more directly with the MR interface itself. They practiced placing and manipulating holograms in the environment, anchoring virtual objects to real-world surfaces such as walls and tables. Participants also experimented with interface elements such as typing on virtual keyboards and navigating basic menus. This allowed them to experience firsthand how MR systems handle spatial interaction, gesture-based input, and environmental tracking—elements that are central to the design of operational MR tools.

The goal of the session was not to evaluate a specific MR system, but to familiarize participants with the possibilities and constraints of MR interaction in general. By exposing officers to a range of core functionalities, the workshop aimed to support more grounded and imaginative feedback in the subsequent user study. In combination with the ethical reflection from the first workshop, this hands-on session enabled participants to critically engage with the idea of integrating MR into law enforcement settings and to reflect on what types of AI-supported MR systems might be both desirable and feasible in their professional domain.

4 User Study and Police Exercise

In this section, we present the user study preceding data collection. The core of the study was a user evaluation conducted through a standard police training exercise for vehicle stops, extended with the use of Microsoft HoloLens headsets (see Figure 4). The objective was to investigate how Mixed Reality (MR) technology might support real-time decision-making, situational awareness, and information delivery during police interventions. This phase of the study followed the preparatory workshops, which ensured participants were familiar with both Trustworthy AI principles and the basic functionalities of MR systems.

Fig. 4. User study: Vehicle stop exercise

The vehicle stop scenario used in this study was a standard part of the police training curriculum, familiar to all participants (see Scenario 1). The exercise was conducted without altering the tactical structure, but with the addition of Microsoft HoloLens headsets. Unlike the preparatory MR workshop, the headsets were turned off during the police exercise and served only to simulate the physical experience of wearing MR equipment. By isolating the MR element to the headset itself and keeping the exercise otherwise unchanged, we aimed to elicit grounded feedback rooted in operational experience rather than software functionality.

The police exercise was conducted as follows:

Scenario 1. *The officers were instructed to conduct a vehicle stop following a standardized police procedure. They operated in two-officer units and were briefed that a wanted suspect was traveling in a vehicle they had located, prompting them to initiate a felony stop. During the exercise, the vehicles were positioned in a static scenario, with approximately two to four car lengths between them. A role player was seated in the driver's seat of the target vehicle, while the patrol unit's vehicle was staged behind it in a tactical position. The exercise coordinator gave the signal to begin the exercise. The suspect in the stopped vehicle was instructed to comply with the officers' verbal commands. If the officers initiated a high-risk exit and ordered the driver out at a distance, the suspect was to attempt to conceal a weapon (a knife) under the vehicle while exiting. The officers received no further directives on how to proceed. Some maintained a safe standoff position, utilizing verbal de-escalation and issuing commands from cover, while others closed distance, conducted a hands-on arrest procedure, and took the driver into custody through physical control techniques. The exercise concluded once the suspect was secured, restrained, and under police control.*

After the exercise, we asked participants to reflect on their experience and articulate what kinds of information should be visualized through MR, when that information would be most useful during the intervention, and how it should be delivered to support safe and effective policing. These reflections formed the basis for the requirement elicitation process described in the following section.

5 Requirement Elicitation

In this section, we outline the methods for data collection and analysis proceeding the user study. To evaluate the officers' experiences and gather system requirements, data collection combined established quantitative and qualitative methods.

5.1 Quantitative Data Collection: Technology Acceptance

Quantitative data was collected through a questionnaire administered after the preparatory workshops and the user study. The items focused on participants' perceived/expected (1) usefulness, (2) ease of use, and (3) trust in AI-based MR-support during police interventions.

There are several models for measuring the acceptance of a technology, but, despite its criticisms (Benbasat and Barki 2007), the model developed by Davis et al. (Davis 1989; Davis et al. 1989), popularly known as Technology Acceptance Model (TAM), is particularly important. It proposes that there are a series of factors (interrelated) that influence the user's attitude towards a new technology, which affects the decision of how and when to use it.

The original approach of the Technology Acceptance Model is based on two main factors:

(1) Perceived Usefulness (PU) defined by Davis as "the degree to which a person believes that using a particular system will make him or his job performance stand out";
(2) Perceived Ease of Use (PEOU) defined by David as "the degree to which a person believes that using a particular system will free him from work effort".

A user's perceptions of these two factors would be determinant in their intention to use a system. However, these factors do not exist in isolation, so it is important to note that:

(1) Beliefs (both PU and PEOU) are influenced by external variables (such as attitude of use, behavioral intention to use, and actual use behavior); and
(2) PEOU has an important effect on the perception of PU.

At this point, the need to extend TAM to address issues related to user perception in systems with an AI component has been pointed out (Vorm and Combs 2022). To meet this need and to improve the evaluation of the acceptance of a technology, we incorporated in our study a tool to evaluate the Degree of Operational Confidence in Artificial systems, the POTDAI tool (Martín-Moncunill et al. 2024), which was designed to be administered as another factor of the TAM and to be complementary to the PU and PEOU factors (Vorm and Combs 2022) (see Figure 5).

Age: _____ Gender:______________ Years of professional service: ___________________________________

HOW DO YOU AGREE WITH THE FOLLOWING QUESTIONS?	Agree			None	Disagree		
	Fully	Quite	Slightly		Slightly	Quite	Fully
1. Using the SmartDigitalCompanion will allow me to complete my tasks faster.							
2. Using the SmartDigitalCompanion will help me complete my tasks as scheduled.							
3. Using the SmartDigitalCompanion will allow me to increase my productivity, completing more work than I could without it.							
4. Using the SmartDigitalCompanion will help me make fewer mistakes and be more efficient.							
5. Using the SmartDigitalCompanion will make doing my job easier.							
6. I find the SmartDigitalCompanion useful for doing my work.							
HOW DO YOU AGREE WITH THE FOLLOWING QUESTIONS?	Agree			None	Disagree		
	Fully	Quite	Slightly		Slightly	Quite	Fully
7. Learning to operate the SmartDigitalCompanion would be easy for me							
8. I think it would be easy for me to get the SmartDigitalCompanion to do what I need.							
9. I believe that the information provided by the SmartDigitalCompanion will be clear, transparent and understandable. I will understand why the system has made a certain decision.							
10. I believe that the interaction with the SmartDigitalCompanion will be flexible and fluid. If I require more information to understand the system's decision I will be able to interact to request it and vice versa.							
11. It will be easy for me to become proficient in operating the SmartDigitalCompanion.							
12. I think the SmartDigitalCompanion will be easy to use.							
HOW DO YOU AGREE WITH THE FOLLOWING QUESTIONS?	Agree			None	Disagree		
	Fully	Quite	Slightly		Slightly	Quite	Fully
13. I believe that the SmartDigitalCompanion will be prone to not give me the most convenient indications.							
14. I think I could easily detect when the SmartDigitalCompanion is not giving me the right indications and redirect the situation.							
15. I think that working with the SmartDigitalCompanion could lead to overconfidence and dependency.							
16. I think I will feel that I will be observed, watched or judged by the SmartDigitalCompanion during the course of my work.							
17. In cases where I am not sure how to proceed, I would normally follow the recommendations provided by the SmartDigitalCompanion.							
18. I believe that the SmartDigitalCompanion will allow me to do my job more safely.							

Fig. 5. Questionnaire: TAM with Trust (adapted from (Martín-Moncunill et al. 2024)).

5.2 Qualitative Data Collection: Pluralistic Walkthrough

To elicit (technical and non-technical) requirements of AI-based MR-systems for police interventions, the pluralistic walkthrough methodology was implemented.

Pluralistic walkthrough (Bias 1994) is a usability testing method used to generate an early evaluation of the design. It consists of a systematic evaluation, to a group of potential user representatives, of a potential design. Participants receive a set of the printed copies, representing the interface and a usage situation, with the dialog boxes they need to perform the given tasks. Evaluators guide users as they perform these simulated tasks printed on paper and provide feedback on these tasks. Documentation or help functions are rarely available at this point, so testers are with the users to inform and answer questions that arise as they do the tasks (so users can do the assigned tasks and designers get feedback to guide future documentation). Then, developers and other team members address concerns and questions have arisen about the interface going forward. In general, usability evaluation with users is recommended in the early stages of system development, and this technique is very well suited to those early stages. It is a very effective and encouraging evaluation method, especially useful for system designers, as they meet users face-to-face in the test environment, easily put themselves in the users' shoes, and are very interested in user feedback (Riihiaho 2002).

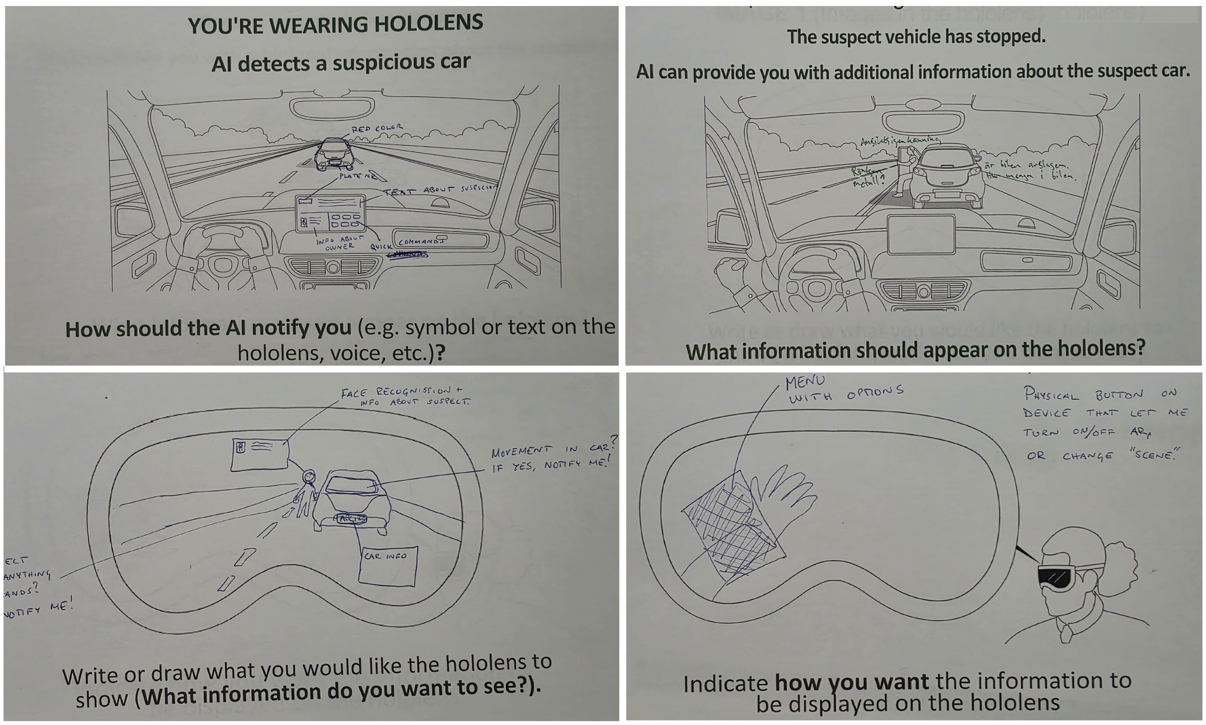


Fig. 6. Pluralistic walkthrough: Illustrations of the traffic stop scenario used by participants to annotate their comments and suggestions.

During the walkthrough, participants were given printed interface mockups and asked to carry out typical tasks they might perform with an AI-based MR system (see Figure 6). As they interacted with the materials, they commented on usability, information delivery, interaction patterns, and interface expectations. Researchers facilitated the sessions, answered clarification questions, and recorded observations and verbal feedback.

5.3 Quantitative Analysis

Statistical tests, such as Mann-Whitney U test or Spearman's correlation, were employed to analyze the collected quantitative data (TAM and POTDAI). This analysis aimed to identify any significant differences in participants' perceptions of the MR device based on demographic factors such as sex, age, and years of service as police officers. The effect size of these significant differences were further explored through the Rosenthal r test.

By understanding participants' perceptions of the MR device's usefulness and ease of use, we can identify areas for improvement and optimization. For example, if certain demographic groups perceive the MR device as less useful or more difficult to use, this information can guide the development of personalized interfaces or training programs to address specific user needs. Furthermore, insights from the TAM analysis can inform the design of the MR technology to enhance transparency, fairness, and accountability, ensuring that it aligns with ethical guidelines for AI deployment in law enforcement contexts. By integrating findings from the TAM analysis with ethical considerations, we can create a more robust and user-centric model for MR technology implementation in police interventions.

5.4 Qualitative Analysis

To analyze participants' feedback on the MR technology (from the pluralistic walkthrough), we applied a Grounded Theory (GT) approach (Oktay 2012), which involves systematically coding and categorizing qualitative data to identify recurring themes and patterns. The analysis aimed to understand officers' perspectives on information sources, modalities, and presentation preferences, supporting the design of processes for information retrieval, personalization, and visualization in MR systems.

To structure our analysis, we employed the General Tactical Explanation Model (GTEM) (Borglund and Hansson 2022) as a theoretical framework. GTEM is a decision-making framework used by the Swedish police to improve situational awareness and facilitate tactical interventions. The model consists of three key stages:

- **Perception**: Detecting and acquiring information from the environment, including risk factors, suspect behavior, and situational elements.
- **Definition**: Interpreting and analyzing the acquired data to assess whether a situation requires intervention.
- **Reaction**: Formulating and executing appropriate tactical responses based on the assessed situation.

By applying GTEM as an analytical lens, the qualitative analysis not only aims to identify technical and non-technical requirements but also to support that the extracted themes align with police decision-making processes.

In the final stage of the analysis, the identified themes and their underlying categories were classified into three risk levels based on Trustworthy AI principles. This helped us illustrate three versions of the envisioned support system, that we call GUARDIAN (low-risk), GUARDIAN+ (medium-risk), and GUARDIAN++ (high-risk). To evaluate the alignment of each system version with ethical guidelines, we applied the European Commission's Trustworthy AI assessment framework, in particular the "Abbreviated Assessment List" released by AI4EU (Dignum, Nieves, et al. 2021), analyzing the implications of each risk category.

6 Quantitative Analysis Results

In this section, we present the quantitative analysis. First, we present the data related to the age, years of service (Table 2) and gender (Table 3) of the agents who participated in the study. Subsequently, Table 4, Table 5 and Table 6 present the results of the technology acceptance model (TAM) test for the Catalonian agents, the Swedish agents and for both groups together. The TAM, originally developed by Davis (1989) and Davis et al. (1989), consists of 2 factors with 6 items (i.e. questions) each (Factor 1: perceived usefulness and Factor 2: perceived ease of use) to which we have added a new factor with another 6 items in this study (Factor 3: confidence of use), see Figure 5. The responses are coded on a likert scale from 1 to 7 according to the degree of agreement of the subject with the statement of the item: 1: agree strongly, 2: agree quite, 3: agree a little, 4: none, 5: disagree a little, 6: disagree quite, 7: disagree strongly (Note: we use inverse scores on items 14, 17, 18 in this new factor for the TAM).

Based on these data, we explored 3 questions: 1) whether there is an influence of age on acceptance and confidence in technology (measured by the TAM); 2) whether years of service influence acceptance and confidence in technology; and 3) whether there are gender differences in acceptance and confidence in technology.

For the analysis of the data, a 95% confidence interval was used, which implied a significance level of 0.05 or less. Given the ordinal nature of the TAM scale, nonparametric analyses were used: Spearman's correlation coefficient and the Mann-Whitney U test were used for between-group contrasts, whose significance was complemented

Table 2. Descriptive data on age and years of service of agents

	n	Minimum	Maximum	Mean	Standard deviation
Age (Catalonia)	20	31	50	40.90	5.180
Years of service (Catalonia)	20	12	19	15.25	2.124
Age (Sweden)	19	30	59	48.47	6.372
Years of service (Sweden)	19	9	37	19.79	7.635
Age (All)	39	30	59	44.59	6.881
Years of service (All)	39	9	37	17.46	5.929

Table 3. Sex of agents

	Catalonia		Sweden		All	
	Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
Male	10	50.0	16	84.2	26	66.7
Female	10	50.0	3	15.8	13	33.3
Total	20	100.0	19	100.0	39	100.0

Table 4. TAM test scores (Original Factors and Factor 3: confidence of use) of the agents of Catalonia

n=19	1		2		3		4		5		6		7		Missing Values	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Factor1.Item1	9	45.0	7	35.0	2	10.0			1	5.0			1	5.0		
Factor1.Item2	6	30.0	9	45.0	3	15.0			1	5.0					1	5.0
Factor1.Item3	9	45.0	6	30.0	5	25.0										
Factor1.Item4	7	35.0	8	40.0	5	25.0										
Factor1.Item5	10	50.0	6	30.0	3	15.0					1	5.0				
Factor1.Item6	13	65.0	6	30.0	1	5.0										
Factor2.Item7	5	25.0	10	50.0	4	20.0	1	5.0								
Factor2.Item8	3	15.0	10	50.0	5	25.0	2	10.0								
Factor2.Item9	6	30.0	7	35.0	4	20.0	1	5.0	2	10.0						
Factor2.Item10	7	35.0	7	35.0	5	25.0			1	5.0						
Factor2.Item11	4	20.0	8	40.0	7	35.0			1	5.0						
Factor2.Item12	5	25.0	9	45.0	3	15.0			3	15.0						
Factor3.Item13			3	15.0	3	15.0	1	5.0	6	30.0	4	20.0	3	15.0		
Factor3.Item14	2	10.0	1	5.0	4	20.0	2	10.0	2	10.0	9	45.0				
Factor3.Item15	4	20.0	6	30.0	5	25.0	2	10.0	2	10.0	1	5.0				
Factor3.Item16	3	15.0	2	10.0	6	30.0	4	20.0	2	10.0	2	10.0	1	5.0		
Factor3.Item17	3	15.0	4	20.0	2	10.0	1	5.0	3	15.0	7	35.0				
Factor3.Item18									5	25.0	5	25.0	10	50.0		

by studying the effect size using Rosenthal's r. The analyses were performed with IBM SPSS Statistics 29.0.2.0 software.

6.1 Influence of Age on the Scores

When studying whether there is a relationship between age and technology acceptance (as represented by TAM item scores), we found:

- In the Catalonia agents, a significant correlation was only found in Item 11 (p-value 0.029) with a Correlation Coefficient of $r = -0.488$ (moderate and inverse correlation).

Table 5. TAM test scores (Original Factors and Factor 3: confidence of use) of the agents of Sweden

n=19	1		2		3		4		5		6		7		Missing Values	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Factor1.Item1	1	5.3	6	31.6	10	52.6	2	10.5								
Factor1.Item2	2	10.5	8	42.1	6	31.6	3	15.8								
Factor1.Item3			10	52.6	6	31.6	2	10.5					1	5.3		
Factor1.Item4	2	10.5	9	47.4	8	42.1										
Factor1.Item5	2	10.5	6	31.6	7	36.8	2	10.5	2	10.5						
Factor1.Item6	6	31.6	10	52.6	3	15.8										
Factor2.Item7	2	10.5	10	52.6	2	10.5	4	21.1			1	5.3				
Factor2.Item8	1	5.3	8	42.1	4	21.1	4	21.1	1	5.3	1	5.3				
Factor2.Item9	1	5.3	5	26.3	7	36.8	5	26.3			1	5.3				
Factor2.Item10	1	5.3	7	36.8	4	21.1	5	26.3	2	10.5						
Factor2.Item11	1	5.3	8	42.1	8	42.1	1	5.3	1	5.3						
Factor2.Item12	1	5.3	11	57.9	3	15.8	3	15.8	1	5.3						
Factor3.Item13					6	31.6	7	36.8	5	26.3	1	5.3				
Factor3.Item14			3	15.8	5	26.3	7	36.8	1	5.3	3	15.8				
Factor3.Item15	1	5.3	4	21.1	6	31.6	3	15.8	3	15.8	2	10.5				
Factor3.Item16	1	5.3	3	15.8	5	26.3	3	15.8	3	15.8	4	21.1				
Factor3.Item17			1	5.3	4	21.1	5	26.3	6	31.6	3	15.8				
Factor3.Item18									12	63.2	3	15.8	4	21.1		

Table 6. TAM test scores (Original Factors and Factor 3: confidence of use) of all the agents

n=39	1		2		3		4		5		6		7		Missing Values	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Factor1.Item1	10	25.6	13	33.3	12	30.8	2	5.1	1	2.6			1	2.6		
Factor1.Item2	8	20.5	17	43.6	9	23.1	3	7.7	1	2.6					1	2.6
Factor1.Item3	9	23.1	16	41.0	11	28.2	2	5.1					1	2.6		
Factor1.Item4	9	23.1	17	43.6	13	33.3										
Factor1.Item5	12	30.8	12	30.8	10	25.6	2	5.1	2	5.1			1	2.6		
Factor1.Item6	19	48.7	16	41.0	4	10.3										
Factor2.Item7	7	17.9	20	51.3	6	15.4	5	12.8			1	2.6				
Factor2.Item8	4	10.3	18	46.2	9	23.1	6	15.4	1	2.6	1	2.6				
Factor2.Item9	7	17.9	12	30.8	11	28.2	6	15.4	2	5.1	1	2.6				
Factor2.Item10	8	20.5	14	35.9	9	23.1	5	12.8	3	7.7						
Factor2.Item11	5	12.8	16	41.0	15	38.5	1	2.6	2	5.1						
Factor2.Item12	6	15.4	20	51.3	6	15.4	3	7.7	4	10.3						
Factor3.Item13			3	7.7	9	23.1	8	20.5	11	28.2	5	12.8	3	7.7		
Factor3.Item14	2	5.1	4	10.3	9	23.1	9	23.1	3	7.7	12	30.8				
Factor3.Item15	5	12.8	10	25.6	11	28.2	5	12.8	5	12.8	3	7.7				
Factor3.Item16	4	10.3	5	12.8	11	28.2	7	17.9	5	12.8	6	15.4	1	2.6		
Factor3.Item17	3	7.7	5	12.8	6	15.4	6	15.4	9	23.1	10	25.6				
Factor3.Item18									17	43.6	8	20.5	14	35.9		

- In the Sweden agents, the following significant correlations were found: Item 3 (p-value 0.033) with a Correlation coefficient of $r = -0.489$ (negative moderate correlation); Item 5 (p-value 0.031) with a Correlation coefficient of $r = -0.495$ (negative moderate correlation); Item 6 (p-value 0.012) with a Correlation coefficient of $r = -0.562$ (negative moderate correlation) and Item 13 (p-value 0.003) with a Correlation coefficient of $r = 0.636$ (good correlation).
- Taking the two groups of agents together only found one significance in Item 4 (p-value 0.044) with a Correlation coefficient of $r = 0.324$ (weak correlation).

6.2 Influence of the Years of Service on the Scores

When studying whether the years of service could influence the approach to technology, we observed the following results:

- In the Catalonia agents, a significant correlation was only found in Item 7 (p-value 0.001) with a Correlation Coefficient of $r = -0.685$ (good inverse correlation).
- No significant correlation between years of service and scores on any item was observed in the Sweden agents.
- Taking the two groups of agents together, significant correlations were found in Item 10 (p-value 0.027) with a Correlation coefficient of $r = 0.354$ (weak correlation) and Item 18 (p-value 0.008) with a Correlation coefficient of $r = -0.417$ (moderate and inverse correlation).

6.3 Comparison of TAM Scores by Gender

We compared the scores of men and women on the items of each factor of the TAM in search of possible significant differences in the approach to technology:

- The following significant differences were found among the Catalonia agents: There are differences by sex in: the factor 2 (perceived ease of use) in: Item 7 ($u = 19.500$, p-value 0.013, with $r : 0.558$ large effect), Item 8 ($u = 25.500$, p-value 0.045, with $r : 0.447$ moderate effect) and Item 11 ($u = 21.000$, p-value 0.020, with $r : 0.520$ large effect); and in the factor 3 (confidence of use) in: Item 16 ($u = 19.000$, p-value 0.017, with $r : 0.534$ large effect) and Item 17 ($u = 22.000$, p-value 0.029, with $r : 0.487$ moderate effect, near to large effect).
- In the Sweden agents, only differences were found in the Item 7 of the factor 2 (perceived ease of use) ($u = 4.500$, p-value 0.018, with $r : 0.379$ moderate effect).
- Taking the two groups of agents together, the following significant differences were observed between the sexes: In factor 1 (perceived usefulness) in Item 5 ($u = 99.500$, p-value 0.031, with $r : 0.344$ moderate effect, near to small effect); in factor 2 (perceived ease of use) in Items 7 ($u = 88.500$, p-value 0.009, with $r : 0.415$, moderate effect) and 11 ($u = 96.500$, p-value 0.021, with $r : 0.370$ moderate effect); and in FACTOR 3 (confidence of use) in Item 15 ($u = 99.000$, p-value 0.033, with $r : 0.341$ moderate effect, near to small effect).

All in all, the quantitative analysis reveals generally positive perceptions of the AI-based MR system across all three evaluated factors: perceived usefulness, ease of use, and confidence of use—the latter introduced as a third factor to capture trust-related dimensions. While age and years of service showed moderate correlations with specific items (particularly among Swedish agents), these demographic variables did not consistently affect overall acceptance. Gender differences were more consistently significant, especially for items related to ease of use and trust. The addition of trust as a measured factor provided valuable insight into users' perceived reliability and confidence in the system. These findings support the importance of addressing trust explicitly in the requirement elicitation of AI-based tools, further explored in the qualitative analysis that follows.

7 Qualitative Analysis Results

In this section, we present the qualitative analysis, resulting in a categorization of information and technology requirements in the studies scenario. The qualitative data obtained from participants' feedback on the MR technology was analyzed using a Grounded Theory (GT) methodology (Oktay 2012). In particular, the approach in (Brännström et al. 2024) was applied, which has shown potential in capturing human perceptions into computational models. This approach involves a systematic examination of qualitative data to generate theories and knowledge models. Through iterative coding and categorization, we identified recurring themes and patterns in participants' responses regarding information sources, modalities, and presentation preferences on the MR device. Let us proceed with presenting each step in the GT process.

7.1 Initial Coding

In the initial coding step of the GT analysis, we systematically examined the qualitative data collected from participants in both the Spain and Sweden user studies. The data consisted of participant quotes and notes regarding their information requirements and preferences for the MR technology during police exercises. Table 7 and Table 8 present data samples along with the corresponding codes. For example, Participant SP01 (Spain) expressed the need to "be able to see people inside the vehicle", leading to codes such as [Person Identification], [Visual Observation]. Similarly, Participant SW06 (Sweden) emphasized "possibility to open more info about the car owner", coded as [Detailed Vehicle Information], [Expandable Data]. In this way, a range of codes emerged, reflecting diverse information requirements for police officers inherent in vehicle stop interventions. These codes encapsulate a broad spectrum of data points, from [suspect identification] and [vehicle registration] to [traffic conditions], [officer experience], [weapons], [explosives], and [hidden objects]. Other codes regard more nuanced aspects, such as [suspicious behavior indicators] and [communication protocols].

These initial codes are the building blocks for our further analysis and categorization. Through iterative coding of participant quotes, we identified recurring themes and patterns related to information and functional requirements for the MR technology.

7.2 Categorization

In the categorization process, the data from the initial coding stage has been organized into categories and subcategories, facilitating a structured understanding of the information requirements in the studied scenario. In the categorization process, we further incorporated insights from the Vehicle stop activity, in which officers mentioned the need for particular information at different stages of the scenario, for the different officer roles; cover and contact. The overall categorization resulting in the following overall categories: [Suspect Information], [Vehicle Information], [Environment Information], [Officer Information], [Suspicious Activity], [Risk Causes], [Crime Information], [Communication & Warning], [Personalization & Modalities], [Sensors & AI Services], [Officer Roles] and [Activity Stages]. Let us summarize the 12 overall categories (see Table 9):

- **Suspect Information:** This category includes a wide range of data crucial for identifying and assessing suspects during vehicle stops. It includes facial recognition, tattoo recognition, and identity verification to confirm the suspect's identity. Additionally, information on criminal records and related crimes provides insight into the suspect's background and potential risk level. Indicators such as passenger detection and suspicious activity recognition help officers assess the situation and take appropriate action.
- **Vehicle Information:** This category focuses on information related to the vehicle involved in the stop. It includes details such as vehicle specifications, status, and license plate expansion. Officers rely on this information to determine the legal status of the vehicle, its location via GPS positioning, and any indicators of suspicious behavior, such as engine status and speed.
- **Environment Information:** This category pertains to data about the surrounding environment during the vehicle stop. It covers factors like traffic and road conditions, rearview data, and the presence of bystanders nearby. Monitoring tools such as environmental sounds and drone overviews provide additional situational awareness to officers, enabling them to assess potential risks and make informed decisions.
- **Officer Information:** This category involves information about the officers themselves, including their experience, role, and physical and mental status. It also includes indicators of officer performance, such as situational awareness and adherence to protocols. Officers rely on this information to assess their readiness and effectiveness in handling the situation.

Table 7. Initial coding from the user study: Spain. Information requirements. (Sample).

Participant	Quote	Codes
SP01	"Be able to see people inside the vehicle."	Person Identification, Visual Observation, Situational Awareness, Communication
SP01	"Be able to have information on the environment."	Environmental Awareness, Situational Context
SP01	"Detecting evidence of crime."	Crime Detection, Evidence Collection
SP01	"Constant communication with the control room."	Communication Tools, Coordination
SP02	"Detect if there are any weapons inside the vehicle."	Weapon Detection, Threat Assessment
SP02	"Provide information about the vehicle (name of the owner, if it is a stolen vehicle)."	Vehicle Information Retrieval, Ownership Details
SP02	"Show the distance to the car in front."	Distance Measurement, Spatial Awareness
SP02	"All the police information you may need to quickly resolve the situation: background information on the person, administrative data, license plate readings..."	Comprehensive Data Access, Rapid Resolution
SP03	"Is it a Stolen vehicle."	Vehicle Theft Detection, Risk Assessment
SP03	"Level of risk."	Risk Evaluation, Threat Perception
SP03	"Detection of persons."	Person Identification, Situational Awareness
SP03	"Inform other patrols."	Inter-Agency Communication, Collaboration
SP04	"AI could have a heat sensor to identify people through objects."	Thermal Imaging, Person Detection
SP04	"Thermal sensor that identifies how many occupants or animals are in the vehicle."	Occupant Detection, Environmental Awareness
SP04	"Detect tattoos or significant features that may provide information."	Biometric Identification, Tattoo Recognition
SP04	"Detect if the plate is modified."	License Plate Verification, Vehicle Status Assessment
SP05	"Passengers number and where they sit."	Passenger Count, Seating Arrangement
SP05	"Extend license plate: vehicle details, vehicle owner, photo of the vehicle owner."	License Plate Expansion, Vehicle Identification
SP05	"Facial recognition of the person in order to be able to compare it with the photo of the owner of the vehicle."	Facial Recognition, Identity Verification
SP06	"Thermal sensor detects 2 persons in the vehicle."	Thermal Imaging, Occupant Detection
SP06	"The suspicious has a possible revolver."	Weapon Identification, Threat Assessment
SP06	"Record the Police action."	Activity Logging, Documentation
SP06	"Voice activation to call for a back-up patrol or other services (ambulance)."	Voice Command, Emergency Response
SP07	"Voice message about plate, administrative data, car model, vehicle owner, color..."	Auditory Alerts, Data Transmission
SP07	"HoloLens vision: passengers number, metal objects, weapons."	Visual Display, Object Detection
SP07	"Communicate with the control room through the HoloLens to keep hands free."	Hands-Free Communication, Operational Efficiency
SP07	"Arrows identifying persons and/or objects of police interest."	Symbolic Indicators, Target Identification
SP08	"License plate and all of information with only one lecture of the plate."	License Plate Retrieval, Data Efficiency
SP08	"Thermic sensor able to detect passengers."	Thermal Imaging, Occupant Detection
SP08	"Facial recognition and compare with police data base."	Facial Recognition, Database Integration
SP08	"Detection of hidden or trapped persons in any place of the vehicle."	Trapped Person Detection, Safety Assessment
SP09	"Information of police interest about the vehicle."	Relevant Vehicle Data, Law Enforcement Needs
SP09	"Car mirror: Information about the behind vehicle."	Rearview Data, Surveillance
SP09	"Ubicate the screen in front of the co-driver. He is going to manipulate."	Display Positioning, User Interaction
SP09	"Drone view."	Aerial Surveillance, Enhanced Perspective
SP10	"Identify license plate, insurance, ITV, theft..."	Vehicle Documentation, Theft Detection
SP10	"Incomming patrol cars, e.g., in the car mirror."	Nearby Patrol Identification, Situational Awareness
SP10	"Drone recording to warn traffic that a police action is underway."	Drone Surveillance, Public Safety
SP10	"Monitoring the environment and location of relevant objects or situations: height of the vehicle, weight, knocks on the vehicle, modifications."	Environmental Monitoring, Object recognition, Vehicle specifications

Table 8. Initial coding from the user studies: Sweden. Information requirements. (Sample).

Participant	Quote	Codes
SW03	"No information until I want it. I want to focus on what is happening. As long as the suspect is moving, I do not want any distractions."	Focus on Immediate Situation, Minimize Distractions
SW03	"Physical button that let me turn on/off AR or change “scene”."	User-Controlled AR, Scene Management
SW05	"Keep the interaction simple."	Simplified Interface
SW05	"AR information should be visible in my left visual field, since I am right eye dominated. This should be an option for each user."	Customizable Visual Display, User Preference
SW06	"Possibility to open more info about the car owner."	Detailed Vehicle Information, Expandable Data
SW06	"Indicate danger with visual information. And indicate danger in what way."	Visual Danger Indicators, Level-based Warning System
SW06	"Indicate danger: (Level 1 danger: ASP general register, more information, Level 2 danger: MR suspect register, Level 3 danger: BR Criminal register) = warning	Hierarchical Danger Classification, Warning System"
SW06	"Simple navigation in the HoloLens controlled with eye movements"	Intuitive HoloLens Navigation, Eye-Controlled Interface
SW07	Symbols, voice, signals	Symbolic Representation, Voice Alerts, Communication Signals
SW07	"Only show relevant information about the car."	Contextual Information Display, Relevance Filtering
SW07	"Prevent intervention if it is not necessary."	Intervention Threshold, Minimize Unnecessary Actions
SW15	"Information about the driver's maneuvering."	Driver Behavior Monitoring, Maneuver Information

- **Suspicious Activity:** This category comprises indicators of suspicious behavior exhibited by suspects during the vehicle stop. It includes actions like door openings, dropping objects, and non-compliance with orders. Detection of weapons, explosives, and other suspicious objects helps officers assess the level of threat and respond accordingly.
- **Risk Causes:** This category covers factors contributing to potential risks during the vehicle stop. It includes traffic conditions, speed limit violations, and the vehicle's history, such as theft or involvement in other crimes. Officers use this information to identify potential threats and take proactive measures to mitigate risks.
- **Crime Information:** This category focuses on information related to specific crimes, such as stolen vehicles, wanted persons, and criminal history. It also includes indicators of criminal activity, such as the presence of explosives or fake documents. Officers rely on this information to identify potential suspects and assess the severity of the situation.
- **Communication & Warning:** This category includes various communication methods used by officers, such as verbal and non-verbal communication, police radio, and tactical communication. It also encompasses warning systems and indicators of danger levels, helping officers coordinate their actions and alert others to potential threats.
- **Officer Roles:** During vehicle stops, officers typically assume the roles of contact and cover to ensure safety and efficiency. The contact officer engages directly with the vehicle occupants, handling communication, requesting documentation, conducting searches, and making decisions based on the gathered information. In contrast, the cover officer focuses on providing security and support, monitoring the scene for potential threats, and being ready to assist the contact officer if necessary. This clear delineation of roles enhances safety, streamlines the process, and ensures effective coordination and communication between officers.
- **Activity Stages:** The activity stages of a vehicle stop outline the sequence of interactions between officers and vehicle occupants.Initially, officers assess the situation from “inside the car”, observing and planning

their approach. This is followed by "outside the car", where officers exit their vehicle to engage with the suspects. From this point, there are two main approaches: "them-to-us", where the vehicle occupants are directed to approach the officers, and "us-to-them", where officers approach the suspect vehicle. Each approach has three stages: "approach", where officers move towards or direct the suspects; "close", where they close the distance; and "contact", where direct interaction occurs. Understanding these stages allows officers to anticipate and respond effectively as the situation develops.

In our analysis of officers' requirements for information during vehicle stops, we have identified additional categories that represent the technical aspects essential for meeting their diverse information needs:

- **Personalization & Modalities:** This category regards tailored warning messages and alerts, alongside a variety of modalities for information dissemination. These modalities encompass features like holographic displays and real-time video feeds. Moreover, it encompasses the fusion of sensory data with AI-driven services, aimed at enhancing officers' perceptual acuity and facilitating real-time data interpretation.
- **Sensors & AI Services:** This category highlights the demand for sensors and AI-driven services designed to optimize monitoring and analysis processes. These technologies span a spectrum of tools including cameras, sound sensors, and object recognition systems. Information acquisition depend on advanced capabilities to seamlessly acquire and process real-time data.

The categorization process has provided a structured understanding of officers' information requirements during vehicle stops, encompassing diverse aspects of information requirement. These categories, along with the additional insights into personalization and sensor & AI services, offer an overview of the demands placed on MR technology to support officers in law enforcement (particularly vehicle stop) scenarios.

7.3 Theory

In the final stage of our thematic analysis process, the individual categories identified in the earlier stages were grouped into overarching themes in line with theoretical framework, GTEM, providing an aggregated understanding of the elicited requirements (see Table 10) as follows:

- **Perception**: This theme encompasses the initial stage of gathering critical data elements essential for effective decision-making during law enforcement operations. It includes the acquisition of Suspect Information, Vehicle Information, Environment Information, and Officer Information.
- **Definition**: In this theme, the focus shifts to the analysis and interpretation of acquired information to evaluate the current scenario effectively. It involves assessing Suspicious Activity, identifying Risk Causes, utilizing Communication & Warning systems, accessing Crime Information.
- **Reaction**: This theme revolves around the formulation and execution of appropriate actions based on the assessment of the situation. It entails deploying Personalization & Modalities, defining Officer Roles, and understanding Activity Stages.

The thematic analysis has organized the identified categories into three overarching themes: *Perception*, *Definition*, and *Reaction*. By structuring the information at multiple levels of abstraction, these themes provide insights into law enforcement decision-making, emphasizing the role of information in enhancing situational awareness and guiding the design of Mixed Reality (MR) technology for police officers.

Perception involves gathering critical data such as suspect identification, vehicle details, environmental conditions, and officer status. Effective MR systems should integrate these diverse sources seamlessly, supported by real-time sensors and AI services. *Definition* focuses on analyzing this data to evaluate ongoing scenarios, detect suspicious activities, and assess risk factors. This requires intelligent data processing and structured presentation to help officers quickly interpret and prioritize information. Finally, *Reaction* involves formulating and executing actions based on this assessment, ensuring clear role definitions, structured communication, and

Table 9. Information Categorization.

Suspect Info	**Vehicle Info**	**Environment Info**	**Officer Info**
Facial Recognition, Tattoo Recognition, Identity, Verification, Vehicle Registration, Criminal Record, Indicate Level 1 danger: ASP general register, Indicate Level 2 danger: MR suspect register, Indicate Level 3 danger: BR Criminal register, Related Crimes, Passenger detection, Suspicion Activity Recognition, Has Suspicious Objects, Has Object in hand, Has Hidden Object, Textual Information (Legal State, Related Crimes), Suspect Physical Status, Suspect Mental Status, Suspect Biographics, Body language	Vehicle specifications, Vehicle status, License Plate Expansion, Vehicle Identification Vehicle stolen Wanted Vehicle Legal state Engine turned on Speed GPS Position	Traffic Conditions, Road Conditions, Rearview Data, Distance to Other Cars, Identification of suitable stop area, Environmental Sounds, Monitoring, Drone overview "Birds view", Bystanders in Proximity, Other cars approaching, Backup Time,	Experience, Role, Biographics, Officer Physical Status, Officer Mental Status, Blind spot Indicators, Tunnel Vision Indicators, Stress Level Indicators, Relaxation Indicators, Situational Awareness, Recognition of officer activity, Dispatch Instructions Adherence, Emergency Response Protocol Adherence, Unnecessary Intervention Indicators
Suspicious Activity	**Risk Causes**	**Crime Info**	**Officer Roles**
Person comes out, Door opening, Dropping of objects, Driver maneuvering, Weapons, Explosives, Hiding objects, Non-compliance to orders, Shouts, Voices, Suspicious objects, Smells, drugs, alcohol	Traffic conditions, Speed limit violations, Car stolen, Car related to other crimes, Crime history of related person, Trapped Person Detection, Weapons, Explosives, Door opening, Person comes out, Backup time long, Non-compliance to orders, Passengers, Tunnel vision, Suspicious behavior, Suspicious objects, Non-compliance to orders, Engine on, People approaching, Other cars approaching, High stress, Activity in the car, Engine starts,	Vehicle stolen, Wanted Vehicle, Vehicle related to other crimes, Wanted person, Legal state, Related person, Owner of suspicious car, Crime history of related person, Explosives, Metal objects, Not following orders, Detect fake documents	Cover, Contact
Communication & Warning	**Personalization & Modalities**	**Sensors & AI Services**	**Activity Stages**
Verbal Communication officers, Non-verbal Communication officers, Tactical Communication, Police Radio Communication, Backup Officer Coordination, Indicate Level 1 danger: ASP general register, Indicate Level 2 danger: MR suspect register, Indicate Level 3 danger: BR Criminal register, Cover Blind Spots	Short Warning Messages, Voice Alerts/Keywords, "Pling" sound for warnings, Text about suspicion, Holographic information, Bounding boxes on Suspects, Bounding boxes on Bystanders, Bounding boxes on Objects, Red color line around car, Red blinking ring around car indicating suspicion, Box following suspect movement, Box indicating nearby bystander, Video feeds, Camera feeds, Social media feeds, X-ray vision, Zoom in on objects, Zoom in on suspect	Worn Camera Thermal camera Drone camera Environment monitoring Sound sensor Scanning documents Odor sensor Geo position Network resource Network resource Web resource Social media Register resource Regnr recognition Material detection Speed Sensor Distance sensor Worn Sensors Worn biometric sensor Eye tracking Camera Content Analysis Object recognition Activity recognition Gesture recognition Voice recognition Face recognition	Inside the car, Outside the car, Approach them-to-us, Approach us-to-them, Close them-to-us, Close us-to-them, Contact them-to-us, Contact us-to-them

Table 10. Themes through the lens of GTEM: Perception, Definition, Reaction.

GTEM:Perception	GTEM:Definition	GTEM:Reaction
Suspect Information	Suspicious Activity	Personalization & Modalities
Vehicle Information	Risk Causes	Officer Roles
Environment Information	Communication & Warning	Activity Stages
Officer Information	Crime Information	Sensors & AI Services
Sensors & AI Services	Sensors & AI Services	

well-coordinated interventions. MR technology should enhance these responses by providing adaptable interfaces, efficient communication tools, and strategic support tailored to dynamic situations.

Given the range of AI-assisted capabilities identified in the thematic analysis, it is crucial to assess their implications for trustworthiness and ethical deployment in high-stakes law enforcement settings. The varying levels of AI involvement—from passive information gathering to automated decision support—raise concerns about human oversight, transparency, and potential risks associated with system recommendations or alerts. To systematically evaluate these concerns, the next section applies the European Commission's Trustworthy AI assessment list to categorize the envisioned systems and analyze their compliance with ethical guidelines.

7.4 Risk Categorization of AI-supported Features

This section builds on the thematic analysis of requirements elicited from police officers, categorizing the requested features based on their potential risk level in critical intervention scenarios. The classification follows a low, medium, and high-risk framework, assessing each feature's impact on situational awareness, decision-making, and officer behavior. By considering the overarching themes identified in the previous analysis, concepts related to Perception, Definition, and Reaction suggest different risk classes in AI-technology. Table 11 presents this categorization, distinguishing between low-risk features, which provide passive support without influencing officer actions, medium-risk features, which involve AI-assisted analysis but still require human interpretation, and high-risk features, which include automated recognition and alerts that may directly affect officer reactions.

To further evaluate the trustworthiness of the envisioned decision-support system for police interventions, we define three possible system versions that align with different risk classifications:

- **GUARDIAN (Low-Risk System)** The system focuses on situational awareness without automating decision-making or influencing officer behavior. It provides basic vehicle information, including specifications, stolen/wanted status, and GPS positioning, helping officers verify relevant details quickly. Environmental awareness is enhanced through traffic and road condition monitoring, proximity detection of bystanders and other vehicles, and backup time estimation to assist with planning safe interventions. Officers can communicate through police radio and receive general environmental monitoring data, such as weather and traffic updates. The system does not provide direct analysis or alerts, ensuring that all assessments remain fully in the hands of the officers.
- **GUARDIAN+ (Medium-Risk System)** Building on GUARDIAN, the system adds moderate AI-assisted analysis to improve decision-making support without removing officer control. The system expands with license plate recognition, criminal record lookups, and passenger detection, allowing officers to access more structured information. It introduces automated tracking of suspicious activities, bounding boxes for suspects, objects, and bystanders, and hidden object detection to assist in recognizing potential threats. Officers receive text-based suspicion warnings and can access drone overviews and external camera feeds for a broader view of the scene. Additionally, monitoring of compliance with police instructions ensures

Table 11. Risk Classification of AI Features in Police Interventions

Low Risk **GTEM:Perception**	**Medium Risk** **GTEM:Definition**	**High Risk** **GTEM:Reaction**
Vehicle specifications	License plate expansion	Facial recognition
Vehicle status (stolen, wanted)	Criminal record lookup	Tattoo recognition
GPS position of vehicles	Passenger detection	Identity verification
Traffic conditions	Suspicious activity recognition	Suspect physical assessment
Road conditions	Detecting hidden objects	Suspect mental assessment
Environmental monitoring	Monitoring compliance with police orders	Detecting weapons/explosives
Distance to other cars	Behavior tracking; Bounding boxes on suspects	Detecting metal objects
Identification of a suitable stop area	Bounding boxes on objects	Door opening alerts; Person exiting alerts
Bystanders in proximity	Bounding boxes on bystanders	Stress level indicators
Backup time estimation	Text-based warnings about suspicion	Tunnel vision indicators
Police radio communication	Drone overview	Worn sensors; Live biometric tracking
	Dispatch compliance logging	Voice alerts suggesting risk
	Camera feeds from external sources	Pling sounds / auditory alerts
	Social media feeds for intelligence	Flashing risk indicators
	Activity recognition	Automated detection of non-compliance to orders
		Personalization in the provision of information according to Officer's roles

that officers are aware of potential defiance. However, all alerts remain non-intrusive and require active checking by the officer rather than immediate response triggers.

- **GUARDIAN++ (High-Risk System)** Building on GUARDIAN+, the system integrates high-risk AI capabilities, introducing biometric recognition and real-time behavior analysis to automate elements of intervention assessment. The system includes facial recognition, tattoo recognition, and identity verification, allowing automatic suspect identification. It further expands suspect state assessments, analyzing physical and mental conditions, stress levels, and tunnel vision indicators, potentially influencing officer decision-making. Real-time weapon and explosives detection, metal object identification, and door opening/person exiting alerts enhance threat detection. Officers receive automated auditory and visual alerts, including voice warnings, pling sounds, and flashing risk indicators, designed to prompt immediate reactions. Automated non-compliance detection and officer role assignment support real-time strategic planning but also increase the risk of system-driven decision-making during critical interventions.

Given these classifications, we assessed the trustworthiness of each system version according to the European Commission's Trustworthy AI guidelines.

7.5 Trustworthy AI Assessment

Using the abbreviated Trustworthy AI assessment list (Dignum, Nieves, et al. 2021), which includes 15 questions on aspects such as system development, stakeholder involvement, and impact, we evaluated the envisioned GUARDIAN system variants. Since these systems are not yet deployed, the assessment focused on the challenges in the involved AI-services for meeting these criteria in the setting of law enforcement. Table 12 present the ethical requirements addressed by each question, along with the assessments (see Figure 7).

Table 12. Trustworthy AI assessment for the GUARDIAN system variants.

Question	GUARDIAN	GUARDIAN+	GUARDIAN++
(Q1): Fundamental Rights	1	1	0
(Q2): Privacy & Data Protection	2	1	0
(Q3): Transparency Rights	1	1	0
(Q4): Accessibility	2	2	1
(Q5): Education & Tutorials	2	1	1
(Q6): Data Management	2	1	0
(Q7): Security	2	2	1
(Q8): Ease to Deactivate/Remove	2	1	0
(Q9): Ease to Access Services Without AI	2	1	0
(Q10): Open-Source Code	2	2	1
(Q11): Ownership	2	2	1
(Q12): Openness Over Data Governance	1	1	0
(Q13): Legislation & Policy	2	2	1
(Q14): Design Impact Assessment	1	1	1
(Q15): Right to Contest/Liability	2	1	0

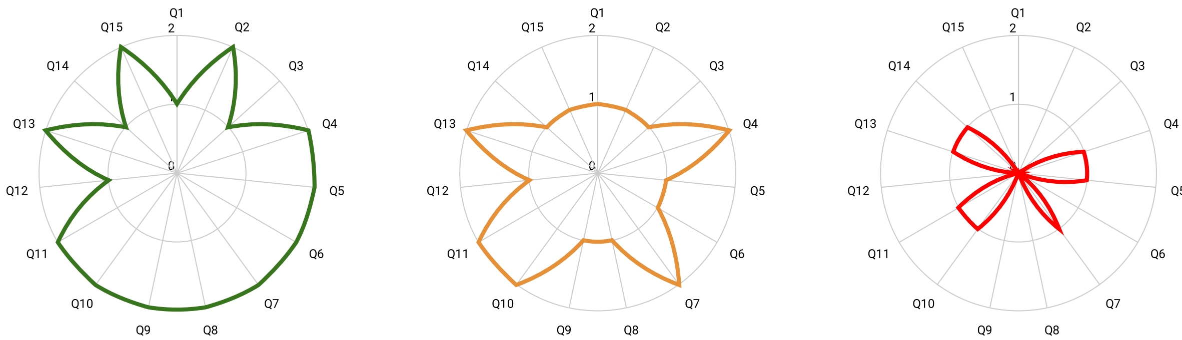
The GUARDIAN system scored well in most areas as it primarily provides situational awareness without AI-driven decision-making, minimal use of sensors and AI-based data interpretation. Fundamental rights (Q1) received a moderate score (1) due to risks from outdated or inaccurate vehicle data, which could lead to unjust stops. Privacy and data protection (Q2) scored high (2) since it collects minimal data and remains GDPR-compliant. However, transparency (Q3) scored lower (1), as the reliability of data sources—such as vehicle status updates and estimated response times—is not always verifiable. Accessibility (Q4) and training (Q5) scored well (2), as the system is widely usable, though officer interpretation errors could still occur. Data minimization (Q6) was rated high (2), as only essential information is collected. Security (Q7) also received a high score (2), due to availability of strong authentication and encryption measures. We consider that the system can be deactivated (Q8) and, as the system does not alarm in real-time, it has less effect on reliance (Q9), both scoring (2). Open-source development (Q10) and ownership (Q11) can be managed, giving a score of (2), but data governance (Q12) received a lower score (1), since not all data sources (e.g., traffic reports) are open or verifiable. Compliance with laws (Q13) was rated high (2), but impact assessment (Q14) was considered challenging, as evaluating potential bias in vehicle data reliability is difficult, especially when it indirectly influences officer decision-making (1). Lastly, officers can retain full decision-making authority (Q15) with a score of (2).

The GUARDIAN+ system builds on the previous version but introduces automated suspect identification, behavior tracking, and criminal record lookups, leading to greater privacy and bias concerns. Fundamental rights

(Q1) scored lower (1), as automated profiling could disproportionately flag certain individuals based on prior records. Privacy (Q2) was also reduced (1), since passenger detection and behavior tracking introduce new data collection risks. Transparency (Q3) remained low (1), as automated recognition systems (e.g., bounding boxes) make data sources harder to verify. Accessibility (Q4) remained high (2), but training (Q5) dropped (1), as officers may misinterpret automated alerts. Data minimization (Q6) was not fully met (1), because additional data, such as criminal records and passenger information, may not always be necessary for situational awareness. Security (Q7) remained strong (2), ensuring controlled access to sensitive data. However, deactivation (Q8) and reliance (Q9) scores dropped (1), since officers may become overly dependent on AI-generated alerts. Open-source development (Q10) can be provided (2). Ownership (Q11) was clear (2), but data governance (Q12) was considered only partially transparent (1), as the system relies on third-party data (e.g., criminal record databases) without clear governance oversight. Legislation and Policy (Q13) must be thoroughly addressed (2), but impact assessment (Q14) proved challenging, as evaluating the fairness of automated criminal record flagging is difficult, particularly regarding false positives and potential profiling biases (1). Officer control (Q15) dropped (1), as AI-based suspect tagging influences decision-making.

The GUARDIAN++ system integrates facial recognition, real-time stress monitoring, and automated non-compliance detection, risk inference and real-time warnings, posing severe risks to privacy, autonomy, and accountability. Fundamental rights (Q1) received the lowest score (0), as biometric tracking directly affects privacy and freedom of movement. Privacy (Q2) was also at its lowest (0), since officers cannot avoid processing sensitive biometric data, making GDPR compliance highly problematic. Transparency (Q3) remained critically low (0), as AI-driven decisions—such as stress level detection and non-compliance alerts—are fully opaque, leaving officers unable to verify their validity. Accessibility (Q4) can possibly remain high (1), though the complexity of AI features may hinder usability for some officers. Training (Q5) is a challenge (1), as AI-generated alerts could cause unavoidable reactions, increasing escalation risks. Data minimization (Q6) dropped to (0), as the system may need to continuously processes biometric data beyond what is strictly necessary. Security (Q7) remained relatively strong (1), but human factors can be a source for breaches. Deactivation (Q8) and reliance (Q9) scores were the lowest (0), as biometric data persists and officers may become entirely dependent on AI risk assessments. Open-source development (Q10) can be challenging (1) given the requirement for advanced AI-services. Hence, ownership (Q11) can be unclear (1), since multiple stakeholders manage different AI components. Data governance (Q12) can be considered opaque (0), given the need for several AI services based on machine learning, making it unclear who controls and has access to biometric data. Legal compliance (Q13) became uncertain (1), as real-time biometric surveillance falls into legal gray areas. Impact assessment (Q14) can have major challenges, as assessing the reliability of real-time stress and non-compliance detection is difficult due to the risk of false positives leading to unjustified interventions (1). Contesting AI-generated risk assessments (Q15) may also be difficult in real-time interventions (0). Overall, GUARDIAN++ presents extreme risks to privacy, transparency, and autonomy.

The assessment highlights that as AI-based services in law enforcement become more advanced, the challenges of ensuring ethical aspects such as fairness, transparency, privacy, and accountability grow. More complex AI-driven tools—such as suspect identification, behavioral tracking, and biometric monitoring—introduce risks of bias, wrongful interventions, and diminished autonomy. Additionally, impact assessments become more complex, particularly when AI influences real-time decision-making through predictive analytics and automated risk evaluations and warnings.



(a) GUARDIAN (b) GUARDIAN+ (c) GUARDIAN++

Fig. 7. Trustworthy AI Assessment of the GUARDIAN system variants. (a) GUARDIAN (low-risk), (b) GUARDIAN+ (medium-risk), and (c) GUARDIAN++ (high-risk). The assessment shows that as AI complexity increases, privacy risks, autonomy reduction, and data opacity escalate. It further finds that GUARDIAN++ exhibits severe trustworthiness issues, while GUARDIAN remains relatively balanced but requires close evaluation regarding non-discrimination (Q1), transparency (Q3), data governance (Q12), and open development (Q14).

8 Discussion and Related Work

Requirements elicitation (Ahmad et al. 2023; Gorton et al. 2023; Maalej et al. 2023; Ruparelia 2010) has long been recognized as a critical phase of the software engineering life cycle, as early design decisions constrain and shape all subsequent development activities (Ruparelia 2010). For artificial intelligence systems, this phase is particularly consequential and challenging (Aqeel and Khan 2025), often due to the strong coupling between requirements, data, learning models, and system behavior over time (De Silva and Alahakoon 2022; Gorton et al. 2023). Prior work on *Ethics by Design* (Brey and Dainow 2024; d'Aquin et al. 2018; Dignum, Baldoni, et al. 2018) argues that ethical and societal values must be integrated into system development from the outset, rather than addressed retrospectively. However, existing ethical guidelines often remain at a high level of abstraction, offering limited support for how such principles should be operationalized during early development stages. Recent research in requirements engineering (Ahmad et al. 2023; Maalej et al. 2023) highlights that traditional elicitation practices are insufficient for responsible AI, and calls for methods that explicitly account for ethical concerns, human-centered values, and socio-technical risks during requirements elicitation. Moreover, studies on responsible AI implementation (Sanderson, Douglas, et al. 2023; Sanderson, Schleiger, et al. 2024) emphasize that ethical principles frequently involve tensions and trade-offs, which must be identified and reasoned about early, before architectural and model-level commitments are fixed. Taken together, this body of work underscores the need for structured requirements elicitation methods tailored to AI systems, capable of embedding ethical considerations into the software engineering life cycle at the point where they can have the greatest impact.

Accordingly, the current study elicited requirements for AI-based tools in law enforcement and examined their ethical implications. The integration of AI-driven technology in policing presents a fundamental tradeoff: on one side, these tools can significantly improve officers' situational awareness and provide critical information to enhance safety and efficiency; on the other, they raise sensitive ethical concerns that demand careful scrutiny. In particular, issues related to non-discrimination, transparency, data governance, and open development stand out in our analysis. These concerns underscore the importance of ensuring that AI-based solutions do not inadvertently perpetuate biases or infringe on individual rights, as well as the need for clear standards and guidelines in designing such systems.

Findings indicate that while officers recognize the substantial benefits of AI-enhanced modalities for improving situational awareness, trust remains a critical concern, especially when AI-generated insights can significantly influence decision-making under high-stress conditions. Participants expressed a preference for systems that provide contextual support without dictating actions, highlighting the importance of maintaining officer autonomy. However, concerns were raised about reliability, potential bias in algorithmic outputs, and the possibility of information overload. These points emphasize the value of transparent AI confidence indicators and mechanisms to override system suggestions when human judgment deems it necessary. Establishing these transparency and override features can serve as foundational elements for building trust in AI-assisted policing.

Our examination of potential system variations—GUARDIAN, GUARDIAN+, and GUARDIAN++—reveals that as AI sophistication increases, new challenges regarding transparency, bias, and officer autonomy also emerge. More advanced AI could offer richer, real-time insights but at the expense of greater complexity and opaque decision pathways. Integrating these increasingly sophisticated systems into operational contexts necessitates robust data management and oversight measures. Rigorous vetting of data sources, routine checks on model performance, and clear guidelines for officer intervention are essential to maintaining system reliability and fairness. In particular, human discretion should remain paramount in high-stakes situations to prevent undue reliance on automated suggestions.

This study is motivated by trends on the use of AI, interactive systems and sensors in Police practices, reviewed in the following sections.

8.1 AI and Law Enforcement Agencies

In today's fast-paced technological landscape, cutting-edge technologies have become ubiquitous across all sectors of society, including Law Enforcement Agencies (LEAs). Among these technologies, AI and extended reality (including virtual reality, augmented reality, and mixed reality) are having a significant impact on LEAs. The emergence of AI has had both positive and negative implications for LEAs, with the potential to assist policing but also the potential for criminal use. For example, (Hayward and Maas 2021) distinguish three types of criminal use of AI:

(1) crime with AI, where the AI is used as a tool,
(2) crime on AI, where the AI is used as an attack surface, and
(3) crime by AI, where the AI is an intermediary.

This paper focus on the potential of AI to support and improve policing. Regarding data management, LEAs use various types of data to feed the computer systems, including directed data (sensor data or criminal data), automated data (e.g. website cookies or passport scanners), or volunteered data (e.g. Google reviews or social media interactions) (Wessels 2023). This data can be used for administrative functions, forensics, license plate recognition, crime mapping, or evidence organization and analysis, among others (Berk 2021). These tasks can be categorized as predictive policing, real-time identification, or investigation (Wessels 2023). In fact, predictive policing is one area where AI has been widely used in law enforcement (Davies and Krame 2023), along with improved situational awareness (E et al. 2020).

8.2 Sensors and Police Practices

Sensors have become a ubiquitous element of modern life, including in the field of policing. The most prevalent sensors employed by law enforcement are those utilized in traffic control, such as license plate readers and speed cameras. Additionally, sensors have also been integrated into officer patrol vehicles, primarily for the purpose of recording video evidence, but also to increase agency accountability and enhance officer-related behaviors, among others (Chapman 2016). Wearable sensor technology is another common type of sensor used by law enforcement agencies, with the body-worn camera being the most popular among them. According to recent statistics, approximately 62% of local police departments in the US have deployed body-worn cameras (Goodison and Brooks 2023). Nevertheless, research findings suggest that the introduction of body-worn cameras may not result in the anticipated reduction in the use of force, contradicting one of the key arguments for their implementation (Ariel et al. 2016). Other examples of wearable sensor technology might include health and fitness trackers and GPS tracking devices. However, recent research indicates that the current state of this wearable sensor technology is inadequate for law enforcement purposes, primarily due to its lack of accuracy and precision (Goodison, Barnum, et al. 2020). Nevertheless, the majority of these technologies are employed by law enforcement agencies with limited knowledge of their potential impacts and without any form of evaluation of their effects (Lum et al. 2019; Strom 2017). Finally, another type of sensors used by law enforcement agencies are those related to criminal investigation. For example, thermography has been used to detect serious environmental crimes (Lega et al. 2014).

8.3 Interactive Systems and Police Practices

Technology has also impacted policing through the use of sensors and extended reality. Research has primarily focused on its application in police training, particularly through serious games. Serious games have been developed for police training in various fields, such as crime scene investigation using mixed reality (Acampora et al. 2023), virtual reality (VR) for traffic accidents (Binsubaih et al. 2006), and virtual reality for investigative interviews (Guimarães et al. 2022). In a systematic literature review (Maneli and Isafiade 2022), the authors found that incorporating light detection and ranging (LiDAR) scanners and immersive technologies, in conjunction with conventional methods, has proven advantageous for crime scene reconstruction. Additionally, serious games have been developed to train police officers in predicting potential terrorist actions (Sormani et al. 2016). Serious games aim to create realistic scenarios and increase situational awareness, enabling trainees to act in a more informed way (Akhgar et al. 2019). One major advantage of VR-based training is the ability for users to engage in risky real-world scenarios within a safe environment. If the VR simulation induces a significant stress response, it can be a more effective training method by fully immersing the user in a realistic three-dimensional environment with minimal risk of injury and user-friendly features (Kamat et al. 2011). Moreover, VR-based training is cost-effective (Farra et al. 2019), requiring less equipment, space, and personnel (e.g., actors for ASD training scenarios), making it a potentially more accessible method for training in high-stress situations like confronting and de-escalating an active shooter or handling high-speed chases.

While research on the use of extended reality in training is well-developed, there have been fewer studies dedicated to mixed-reality applications in real-time police interventions. Some researchers have recognized AR technology as a valuable resource for information exchange in the security domain (Datcu et al. 2015) and enhancing situational awareness by delivering on-the-spot information (Lukosch et al. 2015). However, a study is particularly interesting considering they employed AR in the policing field. The study conducted by (Engelbrecht and Lukosch 2020) investigated the potential use of AR in mobile hotspot policing. The researchers aimed to determine whether AR could assist officers in their work. The study found that the information provided by AR was redundant for experienced officers and that it compromised the attention and situational awareness of officers on-site. Another study demonstrates how AR technology serves as a collaborative tool for crisis management, involving rescue services, police, and military personnel. User feedback indicates that the AR system was well-received, and there was expressed interest in its practical implementation. Additionally, users noted performance advantages when using the AR system compared to traditional tools (Nilsson et al. 2011).

The DARLENE European project is a notable recent example of an ecosystem that integrates innovative AI methods with sensors for activity recognition and pose estimation tasks. It is paired with a wearable AR framework that visualizes inferred results through dynamic content adaptation based on the wearer's stress level and operational context. The concept underwent validation in co-creation workshops with end-users, and experts assessed the decision-making mechanism to enhance LEAs' situational awareness. The assessment outcomes demonstrate that the proposed solution was well-received by the target users in field operations. Furthermore, the situational awareness decision-making mechanism yielded highly satisfactory outcomes. The evaluation of computer vision components showed promising results and identified areas for potential enhancement (Apostolakis et al. 2021).

By considering these observations, our work contributes to the growing body of research at the intersection of AI, extended reality, and public safety, by focusing specifically on the early-stage requirement elicitation process. By grounding this process in both ethical reflection and operational realism, the study highlights the value of participatory methods for surfacing concerns and expectations before implementation begins. As mixed reality technologies evolve and AI capabilities continue to expand, systematic and context-sensitive approaches like the one presented here will be essential for aligning technical innovation with the needs and values of law enforcement professionals.

9 Conclusion

This study introduced a methodology for eliciting ethically grounded and operationally relevant requirements for AI-based mixed reality systems in law enforcement. By combining a Trustworthy AI workshop, immersive MR-based exercises, and complementary quantitative and qualitative analyses, we developed a multi-layered understanding of what officers need from AI systems in high-risk interventions.

The results demonstrate that while officers acknowledge the benefits of AI for situational awareness, they stress the importance of retaining autonomy, understanding system behavior, and being able to override suggestions when necessary. These themes translated into actionable system requirements, which we synthesized into three system variations—*GUARDIAN*, *GUARDIAN+*, and *GUARDIAN++*—representing different levels of complexity and ethical risk. A key insight from this work is the relevance of *Trustworthy AI literacy* during the design and development of AI-based solutions. Supporting participants in understanding ethical principles before eliciting their input enabled more reflective engagement with both technical and social implications. To our knowledge, such a *holistic methodology*—integrating preparatory ethical education and interactive technology exploration before user studies take place—has not been explored in previous AI systems research, let alone within the context of law enforcement. Future work should expand this requirement elicitation framework within the law enforcement domain. This includes broader validation across different units and jurisdictions, longitudinal studies to track evolving needs, and participatory design sessions to refine and prioritize requirements collaboratively. Moreover, this methodology could support the development of domain-specific ethical checklists and risk assessment tools for law enforcement technology procurement and deployment. By embedding ethical alignment early in the requirement elicitation process, we contribute not only to responsible system design but also to enhancing trust, transparency, and operational readiness within policing practice.

Acknowledgments

We are grateful to anonymous referees for their useful comments. We are particularly thankful to all the police instructors and officers who were part of the user study of the research plan reported in this paper.

The authors would like to thank the HumanE AI Network team (https://www.humane-ai.eu), and the institutions participating in the "*To develop a trustworthy AI model for situation awareness by using mixed reality in police interventions*" microproject: Umeå University - Computing Science Department, Umeå University - Police Education Unit, Comet Global Innovation S.L. and Institut de Seguretat Pública de Catalunya -ISPC. The views expressed in this paper are not necessarily those of the HumanE AI Network team neither of the institutions participating in the research plan reported in this paper. This work was partially funded by the Knut and Alice Wallenberg Foundation.

References

R. Abduljabbar, H. Dia, S. Liyanage, and S. A. Bagloee. 2019. "Applications of artificial intelligence in transport: An overview." *Sustainability*, 11, 1, 189.

G. Acampora, P. Trinchese, R. Trinchese, and A. Vitiello. 2023. "A Serious Mixed-Reality Game for Training Police Officers in Tagging Crime Scenes." *Applied Sciences*, 13, 2, 1–14. doi:10.3390/app13021177.

K. Ahmad, M. Abdelrazek, C. Arora, A. A. Baniya, M. Bano, and J. Grundy. 2023. "Requirements engineering framework for human-centered artificial intelligence software systems." *Applied Soft Computing*, 143, 110455.

B. Akhgar, A. Redhead, S. Davey, and J. Saunders. 2019. "Introduction: Serious Games for Law Enforcement Agencies." In: *Serious Games for Enhancing Law Enforcement Agencies*. Ed. by B. Akhgar. Springer Nature Switzerland AG, 1–11.

C. Anglano, M. Canonico, L. Giordano, M. Guazzone, and D. Theseider Dupre. 2021. "User action representation and automated reasoning for the forensic analysis of mobile devices." In: *Proceedings of the 16th International Conference on Availability, Reliability and Security*, 1–7.

O. Aouedi, K. Piamrat, and B. Parrein. 2022. "Intelligent Traffic Management in Next-Generation Networks." *Future Internet*, 14, 2, 44. doi:10.3390/fi14020044.

K. C. Apostolakis, N. Dimitriou, G. Margetis, S. Ntoa, D. Tzovaras, and C. Stephanidis. 2021. "DARLENE – Improving situational awareness of European law enforcement agents through a combination of augmented reality and artificial intelligence solutions." *Open research Europe*, 1, 87–87. doi:10.12688/openreseurope.13715.2.

S. Aqeel and N. A. Khan. 2025. "Challenges and Issues in Requirements Elicitation for Based Systems: A Systematic Literature Review." *Bridging Global Divides for Transnational Higher Education in the AI Era*, 423–446.

B. Ariel, A. Sutherland, D. Henstock, J. Young, P. Drover, J. Sykes, S. Megicks, and R. Henderson. 2016. "Wearing body cameras increases assaults against officers and does not reduce police use of force: Results from a global multi-site experiment." *European journal of criminology*, 13, 6, 744–755.

I. Benbasat and H. Barki. 2007. "Quo vadis TAM?" *Journal of the association for information systems*, 8, 4, 7.

R. Berk. 2021. "Artificial Intelligence, Predictive Policing, and Risk Assessment for Law Enforcement." *Annual Review of Criminology*, 4, 209–237. doi:10.1146/annurev-criminol-051520-012342.

R. G. Bias. 1994. "The pluralistic usability walkthrough: coordinated empathies." In: *Usability inspection methods*, 63–76.

A. Bibal, M. Lognoul, A. de Streel, and B. Frénay. 2021. "Legal requirements on explainability in machine learning." *Artificial Intelligence and Law*, 29, 2, 149–169.

A. Binsubaih, S. Maddock, and D. Romano. 2006. "A serious game for traffic accident investigators." *Interactive Technology and Smart Education*, 3, 4, 329–346. doi:10.1108/17415650680000071.

J. P. Blair, M. H. Martaindale, and W. L. Sandel. 2019. "Peek or push: An examination of two types of room clearing tactics for active shooter event response." *Sage open*, 9, 3, 2158244019871052.

E. A. Borglund and J. Hansson. 2022. "Tactical Police Interventions: Design Challenges for Situational Awareness." In: *ISCRAM*, 1037–1047.

A. Brännström, J. Wester, and J. C. Nieves. 2024. "A formal understanding of computational empathy in interactive agents." *Cognitive Systems Research*, 85, 101203.

P. Brey and B. Dainow. 2024. "Ethics by design for artificial intelligence." *AI and Ethics*, 4, 4, 1265–1277.

L. Cao. 2022. "Ai in finance: challenges, techniques, and opportunities." *ACM Computing Surveys (CSUR)*, 55, 3, 1–38.

B. Chapman. 2016. *Research on the Impact of Technology on Policing Strategy in the 21st Century - Final Report*. RTI International Police Executive Research Forum. https://www.ojp.gov/pdffiles1/nij/grants/251140.pdf.

S. Costantini, G. De Gasperis, and R. Olivieri. 2019. "Digital forensics and investigations meet artificial intelligence." *Annals of Mathematics and Artificial Intelligence*, 86, 1-3, 193–229.

M. d'Aquin, P. Troullinou, N. E. O'Connor, A. Cullen, G. Faller, and L. Holden. 2018. "Towards an" ethics by design" methodology for AI research projects." In: *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, 54–59.

D. Datcu, S. Lukosch, and F. Brazier. 2015. "On the Usability and Effectiveness of Different Interaction Types in Augmented Reality." *International Journal of Human-Computer Interaction*, 31, 3, 193–209. doi:10.1080/10447318.2014.994193.

A. Davies and G. Krame. 2023. "Integrating body-worn cameras, drones, and AI: A framework for enhancing police readiness and response." *Policing: A Journal of Policy and Practice*, 17, 1–13. doi:10.1093/police/paad083.

F. D. Davis. 1989. "Perceived usefulness, perceived ease of use, and user acceptance of information technology." *MIS quarterly*, 319–340.

F. D. Davis, R. P. Bagozzi, and P. R. Warshaw. 1989. "User acceptance of computer technology: A comparison of two theoretical models." *Management science*, 35, 8, 982–1003.

D. De Silva and D. Alahakoon. 2022. "An artificial intelligence life cycle: From conception to production." *Patterns*, 3, 6.

V. Dignum, M. Baldoni, et al.. 2018. "Ethics by design: Necessity or curse?" In: *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, 60–66.

V. Dignum, J. C. Nieves, A. Theodorou, and A. Tubella. 2021. *An abbreviated assessment list to support the responsible development and use of AI.* Technical Report. Department of Computing Sciences, Umeå University.

S. E, K. R, N. E, G. G, D. C, and L. K. 2020. "Trialing Innovative Technologies in Crisis Management—"Airborne and Terrestrial Situational Awareness" as Support Tool in Flood Response." *Applied Sciences*, 10, 11, 3743. doi:10.3390/app10113743.

H. Engelbrecht and S. Lukosch. 2020. "Dangerous or Desirable: Utilizing Augmented Content for Field Policing." *International journal of human-computer interaction*, 36, 15, 1415–1425. doi:10.1080/10447318.2020.1752473.

European Commission. Apr. 2019. *Ethics Guidelines for Trustworthy AI.* https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai. [Online]. (Apr. 2019).

S. L. Farra, M. Gneuhs, E. Hodgson, B. Kawosa, E. T. Miller, A. Simon, N. Timm, and J. Hausfeld. 2019. "Comparative Cost of Virtual Reality Training and Live Exercises for Training Hospital Workers for Evacuation." *Computers, informatics, nursing*, 37, 9, 446–454. doi:10.1097/CIN.0000000000000540.

S. E. Goodison, J. D. Barnum, M. J. Vermeer, D. Woods, S. I. Sitar, S. R. Shelton, and B. A. Jackson. 2020. *Wearable Sensor Technology and Potential Uses Within Law Enforcement.* Tech. rep. RAND - Priority Criminal Justive. https://www.rand.org/pubs/research_reports/RRA108-7.html.

S. E. Goodison and C. Brooks. 2023. "Local police departments, procedures, policies, and technology, 2020–statistical tables." *Washington DC: Officers of Justice Programs, Bureau of Justice Statistics. Journal on October*, 30, 2024.

I. Gorton, F. Khomh, V. Lenarduzzi, C. Menghi, and D. Roman. 2023. "Software Architectures for AI Systems: State of Practice and Challenges." In: *Software Architecture: Research Roadmaps from the Community*. Springer, 25–39.

M. Guimarães, R. Prada, P. A. Santos, J. Dias, C. Soeiro, R. Guerra, C. Steiner-Stanitznig, and A. Molinari. 2022. "ISPO: A Serious Game to train the Interview Skills of Police Officers: ISPO: A Serious Game to train the Interview Skills of Police Officers." *International Journal of Serious Games*, 9, 4, 43–61. doi:10.17083/ijsg.v9i4.514.

R. S. d. S. Guizzardi, G. C. M. Amaral, G. Guizzardi, and J. Mylopoulos. 2020. "Ethical requirements for AI systems." In: *Canadian Conference on Artificial Intelligence*. Springer, 251–256.

K. Hayward and M. Maas. 2021. "Artificial intelligence and crime: A primer for criminologists." *Crime, Media, Culture: An International Journal*, 17, 2, 209–233. doi:10.1177/1741659020917434.

A. Jobin, M. Ienca, and E. Vayena. 2019. "The global landscape of AI ethics guidelines." *Nature Machine Intelligence*, 1, 9, 389–399.

V. R. Kamat, J. C. Martinez, M. Fischer, M. Golparvar-Fard, F. Pena-Mora, and S. Savarese. 2011. "Research in Visualization Techniques for Field Construction: Construction Engineering: Opportunity and Vision for Education, Practice, and Research." *Journal of construction engineering and management*, 137, 10, 853–862.

T. Krafft, K. A. Zweig, and D. Schoch. 2020. "From principles to practice—An interdisciplinary framework to operationalize AI ethics." *Preprint*.

M. Leese. 2021. "Security as Socio-Technical Practice: Predictive Policing and (Non-) Automation." *Swiss Political Science Review*, 27, 1, 150–157.

M. Lega, C. Ferrara, G. Persechino, and P. Bishop. 2014. "Remote sensing in environmental police investigations: aerial platforms and an innovative application of thermography to detect several illegal activities." *Environmental monitoring and assessment*, 186, 8291–8301.

S. Lukosch, M. Billinghurst, L. Alem, and K. Kiyokawa. 2015. "Collaboration in Augmented Reality." *Computer Supported Cooperative Work*, 24, 515–525. doi:10.1007/s10606-015-9239-0.

C. Lum, M. Stoltz, C. S. Koper, and J. A. Scherer. 2019. "Research on body-worn cameras: What we know, what we need to know." *Criminology & public policy*, 18, 1, 93–118.

W. Maalej, Y. D. Pham, and L. Chazette. 2023. "Tailoring requirements engineering for responsible AI." *Computer*, 56, 4, 18–27.

M. A. Maneli and O. E. Isafiade. 2022. "3D Forensic Crime Scene Reconstruction Involving Immersive Technology: A Systematic Literature Review." *IEEE access*, 10, 88821–88857. doi:10.1109/ACCESS.2022.3199437.

N. Marangunić and A. Granić. 2015. "Technology acceptance model: a literature review from 1986 to 2013." *Universal access in the information society*, 14, 81–95.

D. Martín-Moncunill, E. G. Laredo, and J. C. Nieves. 2024. "POTDAI: A Tool to Evaluate the Perceived Operational Trust Degree in Artificial Intelligence Systems." *IEEE Access*.

S. Nilsson, B. J. E. Johansson, and A. Jonsson. 2011. "Cross-Organizational Collaboration Supported by Augmented Reality." *IEEE Transactions on Visualization and Computer Graphics*, 17, 10, 1380–1392. doi:10.1109/TVCG.2010.249.

J. S. Oktay. 2012. *Grounded theory*. Pocket Guide to Social Work Re.

J. Perez-Cerrolaza et al.. 2024. "Artificial intelligence for safety-critical systems in industrial and transportation domains: A survey." *ACM Computing Surveys*, 56, 7, 1–40.

K. Renaud, I. Bongiovanni, S. Wilford, and A. Irons. 2021. "PRECEPT-4-Justice: A bias-neutralising framework for digital forensics investigations." *Science & Justice*, 61, 5, 477–492.

S. Riihiaho. 2002. "The pluralistic usability walk-through method." *Ergonomics in Design*, 10, 3, 23–27.

S. Rokhsaritalemi, A. Sadeghi-Niaraki, and S.-M. Choi. 2020. "A review on mixed reality: Current trends, challenges and prospects." *Applied Sciences*, 10, 2, 636.

N. B. Ruparelia. 2010. "Software development lifecycle models." *ACM SIGSOFT Software Engineering Notes*, 35, 3, 8–13.

C. Sanderson, D. Douglas, and Q. Lu. 2023. "Implementing responsible AI: Tensions and trade-offs between ethics aspects." In: *2023 International joint conference on neural networks (IJCNN)*. IEEE, 1–7.

C. Sanderson, E. Schleiger, D. Douglas, P. Kuhnert, and Q. Lu. 2024. "Resolving ethics trade-offs in implementing responsible AI." In: *2024 IEEE Conference on Artificial Intelligence (CAI)*. IEEE, 1208–1213.

M. Y. Shaheen. 2021. "Applications of Artificial Intelligence (AI) in healthcare: A review." *ScienceOpen Preprints*.

R. Sormani, J. Soldatos, S. Vassilaras, G. Kioumourtzis, G. Leventakis, I. Giordani, and F. Tisato. 2016. "A serious game empowering the prediction of potential terrorist actions." *Journal of Policing, Intelligence and Counter Terrorism*, 11, 1, 30–48. doi:10.1080/18335330.2016.1161222.

A. F. de Sousa Silva, G. Ramos Sousa Silva, and E. D. Canedo. 2022. "Requirements Elicitation Techniques and Tools in the Context of Artificial Intelligence." In: *Brazilian Conference on Intelligent Systems (BRACIS)*. Springer, 15–29.

M. Speicher, B. D. Hall, and M. Nebeling. 2019. "What is mixed reality?" In: *Proceedings of the 2019 CHI conference on human factors in computing systems*, 1–15.

K. Strom. 2017. "Research on the impact of technology on policing strategy in the 21st century, final report." *Washington DC: US Department of Justice*, 1–11.

E. S. Vorm and D. J. Combs. 2022. "Integrating transparency, trust, and acceptance: The intelligent systems technology acceptance model (ISTAM)." *International Journal of Human–Computer Interaction*, 38, 18-20, 1828–1845.

J. Vossers, A. Brännström, E. Borglund, J. Hansson, and J. C. Nieves. 2024. "Human-aware planning for situational awareness in indoor police interventions." In: *HHAI 2024, The third International Conference on Hybrid Human-Artificial Intelligence, Malmö, Sweden, June 10-14, 2024.* IOS Press, 325–334.

M. Wessels. 2023. "Algorithmic policing accountability: eight sociotechnical challenges." *Policing and Society*, 1–15. doi:10.1080/10439463.2023.2241965.

J. B. Woods. 2018. "Policing, danger narratives, and routine traffic stops." *Mich. L. Rev.*, 117, 635.

Received 09 June 2025; accepted 21 January 2026.