AP4AI
Accountability Principles for AI

Accountability Principles for Artificial Intelligence (AP4AI) in the Internal Security Domain

International Citizen Consultation on AI Accountability in Policing

Accountability Principles for Artificial Intelligence (AP4AI) in the Internal Security Domain
International Citizen Consultation on AI Accountability in Policing

October 2023

Coordinated by:

- CENTRIC (Centre of Excellence in Terrorism, Resilience, Intelligence and Organised Crime Research)
- Europol Innovation Lab

Supporting partners:

- Eurojust
- EUAA (European Union Agency for Asylum)
- CEPOL (European Union Agency for Law Enforcement Training)

Disclaimer: This report presents the results of the AP4AI citizen consultation. The consultation summarises views from 6,674 citizens in 30 countries. The combination of quantitative and qualitative findings offers an in-depth reflection by citizens on the AP4AI Principles and Framework, as well as insights into proposed implementation mechanisms. The AP4AI Project is jointly conducted by CENTRIC and Europol and supported by Eurojust, EUAA, and CEPOL, with advice and contributions by the EU Agency for Fundamental Rights (FRA), in the framework of the EU Innovation Hub for Internal Security. The research outcomes, opinions, critical reflections, conclusions, and recommendations stated in this report do not necessarily reflect the views of CENTRIC, Europol, FRA, CEPOL, EUAA, or Eurojust. The project received ethics approval from the university ethics board of Sheffield Hallam University, where CENTRIC is located as the academic lead of the AP4AI Project.

Authors:

- P. Saskia Bayerl, Marcel Obst, Babak Akhgar, *CENTRIC*

Suggested citation:

- Bayerl, P.S., Obst, M., & Akhgar, B. (2023). *International Citizen Consultation on AI Accountability in Policing. AP4AI Report 3*.

Acknowledgement
The AP4AI Project would like to express its sincere appreciation to the large number of citizens who have engaged with the AP4AI Project and generously provided us with their valuable insights.

FOREWORD

AI Accountability refers to the idea that those who use AI should be responsible for its outcomes and impacts and should further be able to explain and justify their decisions to relevant stakeholders, such as the public, courts, regulators, and the victims or perpetrators of crimes. AI Accountability also implies that there should be mechanisms to monitor, audit, review, and correct the performance and behaviour of AI systems and to provide remedies or redress for any harms or errors that may occur.

Providing Law Enforcement Agencies (LEAs) and other stakeholders in the internal security domain with effective mechanisms to implement, assess, demonstrate, and improve AI Accountability is a core ambition of our AP4AI project. The successful implementation of AI Accountability can help foster public trust, ensure legal compliance, and enhance AI governance. Achieving AI Accountability in practice is not a simple or straightforward task. It requires collaboration and coordination among various stakeholders, such as policymakers, legislators, police officers, developers, regulators, researchers, civil society groups, and citizens. It also requires a clear understanding of how AI works, what it can and cannot do, and what its limitations and uncertainties are. Therefore, AI Accountability is not only a technical or legal issue but also a social and political one. The AP4AI approach is thus closely aligned with the spirit and requirements of the proposed European AI Act (AIA).

This report presents the results of the large-scale citizen consultation AP4AI conducted to understand public perspectives about AI use and AI Accountability. The consultation is driven by the acknowledgement of citizens as core partners in all AI Accountability efforts, following Sir Robert Peel's tenet that it is incumbent on police *'to recognise always that the power of the police to fulfil their functions and duties is dependent on public approval of their existence, actions and behaviour, and on their ability to secure and maintain public respect.'*

The findings in this report can help inform policymakers, LEAs, and government sectors about citizens' views on the perceived benefits and risks of AI and what safeguards and oversight should be in place to build and retain public trust. The outcomes can further support governmental policies and objectives on citizen participation in safety and security.

I am grateful to the joint coordinator of the AP4AI project, Europol Innovation Lab. I am also grateful to the CENTRIC Head of Research, Prof. Saskia Bayerl, for leading this important work.

Prof. Babak Akhgar OBE
Director of CENTRIC

EXECUTIVE SUMMARY

This report offers insights from a public consultation with citizens in 30 countries, including all 27 EU Member States. The aim of this extensive citizen consultation was twofold:

1. To investigate and understand public views about AI use by police forces across a broad diversity of communities, jurisdictions, and policing contexts, and
2. To co-shape the development of AI Accountability mechanisms, the AP4AI project developes together with the law enforcement and justice sector.

The consultation results provide important insights into **public reactions** to AI use by police as well as the **concrete mechanisms** citizens request and expect to ensure AI Accountability. It further outlines specific **safeguards** citizens propose to retain citizens' trust in security actors and their use of AI. Our report furthermore offers crucial **validation for the AI Accountability Principles** put forward by the AP4AI project, including citizens own ideas about concrete implementation mechanisms.

Overall, our citizen consultation demonstrates that citizen views about AI use by police are highly varied but also largely supportive for clearly prescribed purposes. At the same time, our findings also indicate that AI deployments by police need to be accompanied by meaningful mechanisms that ensure AI Accountability, trust, engagement, and legitimacy.

This report provides a detailed picture of the very concrete expectations the public holds about AI use by the police—and specifically how AI use should be regulated and how police should be held accountable. The results allow us to formulate **concrete recommendations**, which are detailed at the end of this report.

The AP4AI citizen consultation was conducted in 27 EU Member States, the UK, the USA, and Australia, with a total of 6,647 citizens. The sample recruitment was done by the international panel provider Qualtrics.

OUTLINE OF RESULTS

The consultation demonstrates **considerable support for AI deployments by police**: across all 30 countries, 66.7% agreed or strongly agreed that AI can greatly profit society, compared to 8.3% who disagreed or strongly disagreed. Even higher was the approval for specific application areas: 87.6% agreed or strongly agreed that **AI should be used for the protection of children and vulnerable groups**; 83.0% agreed or strongly agreed that **AI should be used to detect criminals and criminal organisations**; and nearly three in four participants agreed that **AI should be used to predict crimes before they happen** (74.3%).

These findings suggest that large parts of the public find considerable value in the use of AI by police forces if it aims to protect vulnerable groups and society in a meaningful way.

Credible efforts by police forces provide reassurance for citizens. The more police were perceived to make sufficient efforts to avoid negative consequences and show respect for citizens in their activities, the more positive were attitudes towards AI use by police, and the lower the concerns about privacy and negative consequences. These results speak for the need to put adequate mechanisms in place that give citizens trust in the police forces' correct AI use.

Yet, **only a third of participants considered accountability mechanisms, as they currently exist, to be adequate**. 25.8% perceived them as too weak, while 8.1% rated them as too restrictive. At the same time, a substantial part of the **public seems to lack sufficient information about existing mechanisms** to make an informed judgement (34.2% indicating to 'don't know'). AI expertise clearly affected the perception of accountability mechanisms: people who reported having good AI knowledge or being AI experts were more likely to perceive current mechanisms as 'just right' or as 'too restrictive'. In contrast, people who perceived themselves to have no or little knowledge about AI were most often unaware of accountability mechanisms. Our consultation is thus **a clear call to improve public information and education**.

Further, only half of the participants (50.7%) reported that police in their country make sufficient **efforts to avoid the negative consequences of AI**. Further, 61.1% of participants found that **police lack respect for citizens' rights** in their activities, while nearly half of the participants noted **concerns about police using AI to monitor either their online information** (48.7%) or their offline activities (46.7%). Concrete concerns about **potential negative effects** on themselves due to police decisions based on AI were still noted by 37.2% of participants (neutral: 33.0; no fears: 29.8%).

The second part of the consultation aimed to capture whether citizens experience an AI Accountability framework, as proposed by the AP4AI Project, as a viable approach to safeguarding appropriate AI use by police. The consultation offers **strong validation for the AP4AI Framework overall and its AP4AI Principles**. Firstly, we found strong general approval for an overarching AI Accountability Framework (81.8% rated it important or very important, 15.1% neutral, and 3.1% not important).

More specifically, participants felt strongly that **police need to be held accountable for their AI use**: 90.0% expect police to be held accountable for the way they use AI, and 88.3% for the consequences of their AI use. AI Accountability thus emerges as an approach that is expected by society, in that citizens demand strong mechanisms as well as believable reassurances that police deploy AI in an appropriate way.

Considering the AP4AI Principles themselves, all 12 AP4AI Principles were rated as important or very important. Legality, Conduct and Explainability emerged as the most important principles. However, variations amongst the 12 principles were very minor, **validating the relevance of the principle set** as a meaningful foundation for an AI Accountability Framework.

Courts, police, and governments were identified as the **main actors responsible for monitoring** the appropriate police use of AI and for **enforcing sanctions** in cases where AI use is proven to be inappropriate. In contrast, the desire for citizen involvement was limited, both in the form of direct citizen participation and for citizen participation through representatives.

Further, the majority of participants across all 30 countries preferred **regulation of AI use by police within their own country** (56.2%). For EU citizens, this was followed by a preference for regulation by the EU (39.0%). Other options were chosen less frequently: regulation on a global level by 27.6% of participants and regulation by international organisations such as the UN by 21.6%.

CONTENTS

PURPOSE OF THE CITIZEN CONSULTATION

Citizen perspectives about the use of Artificial Intelligence (AI) for policing and law enforcement purposes are complex and diverse.[1,2] This diversity is as much a truism as it is a call to action: firstly, to effectively map and understand this complexity; secondly, to develop mechanisms that can give weight to the variety of expectations and requirements across citizen groups.

This report aims to support these ambitions by mapping out findings from a citizen consultation about AI use by law enforcement agencies (LEAs).

The citizen consultation was conducted in 30 countries, covering the 27 EU Member States, the UK, USA, and Australia. The broad remit of the consultation was decided to allow us access to a highly diverse set of publics in different constituencies, national levels of AI deployments by law enforcement agencies, and regulatory AI regimes.

Tapping into the diversity of public expectations provides a pathway to understanding broadly shared apprehensions and expectations of the public about AI use for security purposes across many cultural, social, and political contexts. At the same time, the extensive dataset also enables us to identify minority voices, which we consider equally important to give expression to in order to acknowledge perspectives, concerns, and expectations that may go unrecognised in many discussions of AI use by police forces.

Drawing on a highly diverse sample of citizens, it was also vital for us to determine which mechanisms can support public confidence that AI is designed and used in an appropriate way. The consultation was thus specifically shaped to collect citizen proposals for mechanisms to ensure AI Accountability. The data further allowed us to identify proposals for safeguards that support the building and/or retention of citizens' trust in security actors.

In this context, we also validated the 12 AP4AI Principles, which we developed empirically through expert consultations and are detailed in previous reports (Akhgar et al., 2022a and 2022b).

The citizen consultation is part of ongoing consultation efforts by the AP4AI project with subject matter experts that underpin the development of AP4AI's conceptual and practical tools to support the law enforcement and justice sectors.[3] AP4AI recognises citizens as subject matter experts in the same way as representatives of police forces, industry, academia, or policymakers, i.e., as experts in the domains that matter to them: security and the protection of their private lives. Not only are members of the public directly affected by AI deployments by security practitioners, but citizens are a core stakeholder in AI Accountability in the security domain.

Civil participation is invaluable in ensuring the public has an input on the impacts of AI use within their communities.[4] Engagement with citizens also reacts to requests that 'policies should prioritise public participation as a core policy goal.'[5] For us, citizen consultation is thus a core element to achieving AP4AI's ambition of creating practical mechanisms and tools that directly and meaningfully support AI Accountability.

ACCOUNTABILITY PRINCIPLES FOR AI: AN INTRODUCTION TO THE AP4AI PROJECT

AP4AI develops solutions that support security and justice practitioners worldwide in capitalising on the opportunities of AI while demonstrating and safeguarding comprehensive accountability of their AI use towards society. AP4AI does so by providing:

1. The **AP4AI Framework for AI Accountability** for Policing, Security, and Justice which offers comprehensive, principle-based guidance. The AP4AI Framework guides JHA actors, AI suppliers, oversight bodies, policymakers, and society in how to implement, regulate, and/or assess AI Accountability, either proactively (e.g., to guide design or procurement decisions) or reactively (e.g., when questioned by citizens or in courts).

2. The **AP4AI self-assessment tool** which is a hands-on software solution for practitioners to assess their specific AI designs and/or deployments for adherence to AI Accountability. The self-assessment covers disparate application areas (policing, border management, public space protection, etc.) and the full lifecycle of an AI capability (from its design or procurement to its deployment, modification, and retirement). The AP4AI tool will also included assessments against the requirements of the EU AI Act and GDPR.

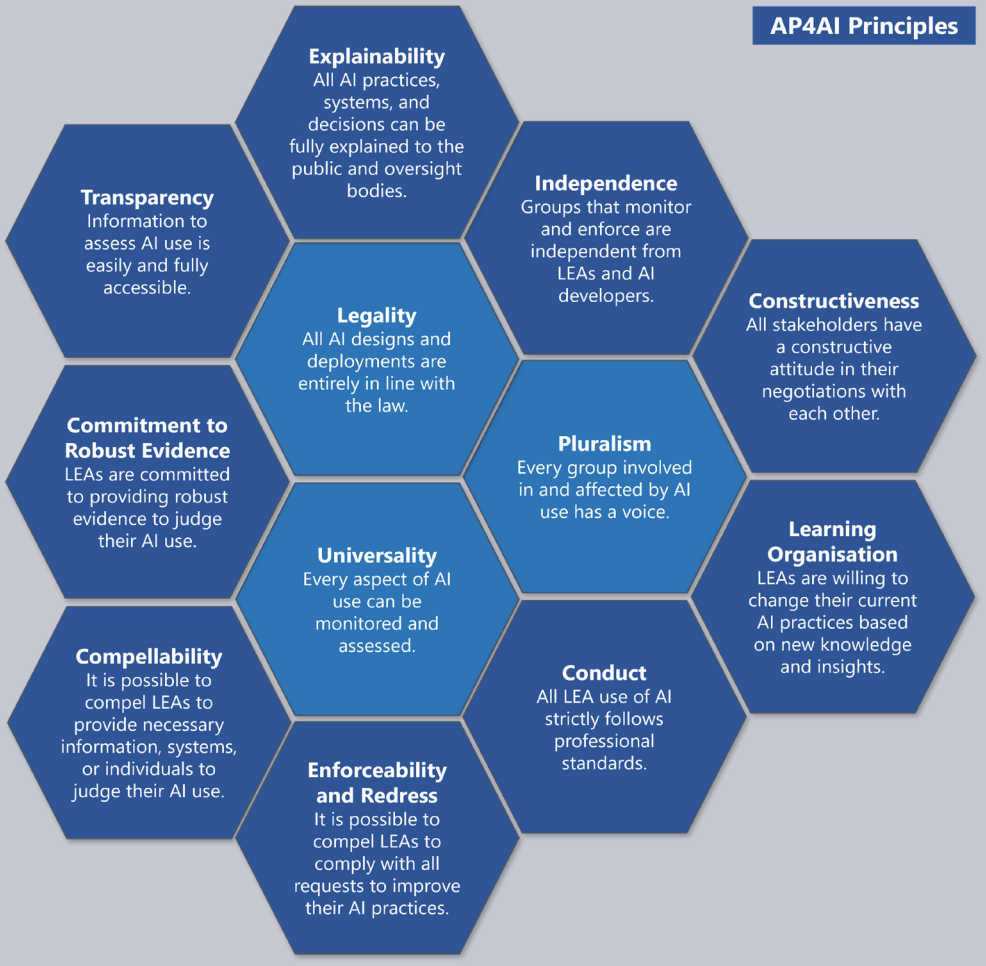
The core innovation of AP4AI is in defining the form and nature of AI Accountability in the security and justice domain, which has so far been missing and thus inaccessible to policy and practical efforts. [6,7,8]

AP4AI uses the following definition of AI Accountability as:

Accountability with respect to the design, purchase, and deployment of AI tools and systems, their outcomes, and their impacts.

Overall definitions of accountability are not useful for judging how well organisations "do" AI Accountability. Rather, AI Accountability needs to be assessable to be meaningful.

AP4AI conceptualises AI Accountability through 12 Principles, developed in extensive expert consultations.



The 12 Principles have then been translated into concrete questions to capture the degree to which AI systems, activities, and procedures adhere to each of the principles' requirements. In this way, AI Accountability becomes directly assessable and demonstrable for the full AI deployment landscape.

At the same time, the overarching principles are 'universal', which ensures that AP4AI can be applied globally while being adaptive across domains, regulatory requirements, national contexts, technical specifications, and further AI innovations.

For further information about the project, its objectives, products, and participating organisations, please consult the project website: www.ap4ai.eu. A comprehensive description of the AP4AI Framework and its methodology is provided in the two reports "AP4AI Framework Blueprint"[9] and "AP4AI Summary Report on Expert Consultations"[10] (both available for download on the AP4AI website).

CONSULTATION METHODOLOGY

ASPECTS ADDRESSED IN THE CITIZEN CONSULTATION

The consultation investigated the following four aspects:

- General attitudes towards AI use by police
- Reactions to the initial set of AP4AI Principles
- Relevance of AI Accountability and Recommendations for Accountability Mechanisms
- Regulation level and actors responsible for AI Accountability

The consultation also provided space for general comments and suggestions. The questions used a mixed format of Likert-scale items (answers ranging from 1–5) and open questions that allowed for free-text answers.

Item formulation: AP4AI as a project addresses the internal security domain. However, this area and the meaning of this term are highly abstract. Instead, the survey items and explanations used 'police' as a reference point, as citizens are likely to have concrete views about police as the most visible representative of security practitioners in their country. The findings of the citizen consultation should thus primarily be interpreted in the context of policing applications of AI.

PARTICIPANTS

The citizen consultation was conducted with adult participants (18 years of age or older) from the general population. To obtain a diverse sample, the consultation covered 30 countries: the 27 EU Member States, the UK, USA, and Australia. Recruitment was intentionally broad, i.e., we did not deploy any preferential recruitment for or exclusion of specific demographic groups, professions, etc. However, gender and age were stratified to be proportional to each country's demographic profile. Further, the sampling ensured proportional regional representation in France, Germany, Italy, Spain, the UK, USA, and Australia. Country samples ranged from 109 (Luxembourg) to 319 (France), for a total of 6,647 participants across all 30 countries.

The sample characteristics demonstrate that the citizen consultation managed to engage a varied set of participants, as intended. Approx. 53% identified as female, 47% as male, 0.5% as non-binary, and 0.1% as other (0.1% preferred not to answer). The sample presents a good spread across educational levels and further includes citizens self-describing as ethnic minorities, as well as citizens with personal experiences of crime and participants working in a security-related profession. Table 1 provides a detailed overview of participant characteristics.

On average, the self-ascribed knowledge about AI was reported as moderate (m = 2.77 on a scale from 1:no knowledge to 5:expert; SD = .95), while the average expertise about AI use by police was reported as limited (m = 2.27 on a scale from 1:no knowledge to 5:expert; SD = .99).

Table 1. Participant characteristics

Sample Size	Gender Distribution*	Age Distribution*	Highest Education	Security-related work	Ethnic Minority	Crime Victim
6,647	Female: 52.6% Male: 46.8% Non-binary: 0.5% Other: 0.1% PNTS**: 0.1%	18-24: 9.0% 25-34: 17.5% 35-44: 19.7% 45-54: 19.7% 55+: 34.1%	No formal education: 0.2% Primary school: 4.0% Secondary school: 39.8% Bachelor or master degree: 36.3% PhD: 2.0% Professional degree: 13.2% Other: 4.2% PNTS: 0.4%	No: 84.5% Yes: 7.8% No work history: 6.6% PNTS: 1.2%	Yes: 12.3% No: 85.8% PNTS: 1.9%	Yes: 37.7% No: 60.8% PNTS: 1.5%

* Pre-determined quotas used to reflect population characteristics in each country; **PNTS: answer option 'prefer not to say'

SURVEY PROVISION

Due to the scale of the engagement, the citizen consultation was conducted as an online survey. The surveys were presented in the respective country's language to ensure that participants could answer questions without language barriers. The translation of the survey from English into the country languages and the recruitment of samples for all countries was organised by the panel provider Qualtrics.

DATA ANALYSIS

The analysis of the quantitative data was conducted on weighted sample sizes, with weighting according to population size to improve proportional representation in the results. The population sizes for each country were taken from the United Nations Data Portal Population Division.[11] Subgroup analyses for demographics were conducted only if sufficient data for a category was available to avoid potentially biased interpretations about demographic groups based on very small participant numbers. For this reason, gender comparisons were only conducted for *women* versus *men*, not for the categories of *non-binary* and *other*.

Open answers were coded thematically by clustering verbatim answers that represented the same idea or theme into the same thematic category (e.g., 'type of AI benefit', 'privacy concerns', 'legal mechanism').

The percentage of participants providing free-text answers varied per survey section: 50% of the full sample provided input for AI Accountability mechanisms, while only 0.4% named additional stakeholders that should be responsible for redress next to the ones mentioned in the quantitative items. About 28% of the participants used the opportunity to provide a comment at the end of the survey.

Participants who were willing to provide additional information often held strong feelings about AI and/or police, which at times caused qualitative answers to diverge in tone or content from the larger picture that emerged from the quantitative data. These answers are thus not necessarily 'representative' for the majority or shared views. They do, however, grant access to concrete examples and expressions of citizens, which we consider important to showcase the type of raw data we received in the citizen consultation, as well as to showcase the high degree of variation and complexity in citizen reactions towards AI use by police. We report open answers in all areas as an opportunity to give voice to these concrete thoughts, expectations, concerns, and hopes of citizens when confronted with AI use by police.

ETHICS

The study received ethics approval from Sheffield Hallam University, the home institution of CENTRIC, which leads the empirical activities in AP4AI. Moreover, all participants were required to give their informed consent before starting the survey. Participants who did not consent were unable to continue with the consultation.

NOTES ON THE SCOPE OF THE CONSULTATION

Our consultation focused explicitly on police to provide participants with a clear and well-known reference point for their answers. This means that other areas of AI use in the internal security domain are not directly covered by the current consultation and would profit from explorations targeted at these specific areas.

While being international in setup, we consciously refrain from discussing detailed country comparisons. Differences among countries were often quite small. Presenting statistical test results risks overstating these small differences, which we consider problematic.

We further draw attention to the fact that younger generations (18–24 years) are underrepresented, with minors (e.g., individuals younger than 18 years) excluded from participation. The perspective of youth is thus not captured and would need further exploration in other projects. Furthermore, while the survey collected information about participants' backgrounds, such as gender, age group, self-description as ethnic minorities, and having been a victim of crime in the past, there are many other aspects that may impact attitudes towards AI (socio-economic status, sexual orientation, being a parent, having a history of migration, etc). Collecting them all was beyond the scope of this consultation, and we invite further consultations to explore additional aspects.

FINDINGS

CONSIDERABLE SUPPORT OF AI USE FOR SPECIFIC PURPOSES

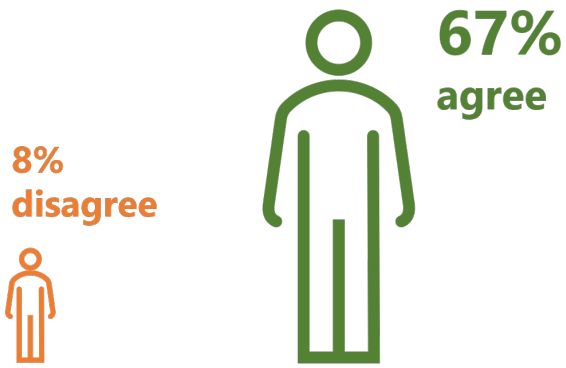
The data shows considerable support for AI deployments by the police: 66.7% agreed or strongly agreed that **AI can greatly profit** society compared to 8.3% that disagreed or strongly disagreed (25.0% neutral).

Even higher was the **approval for specific application areas**: 87.6% agreed or strongly agreed that AI should be used for the protection of children and vulnerable groups; 83.0% agreed or strongly agreed that AI should be used to detect criminals and criminal organisations; and nearly three in four participants (73.4%) agreed that AI should be used to predict crimes before they happen.

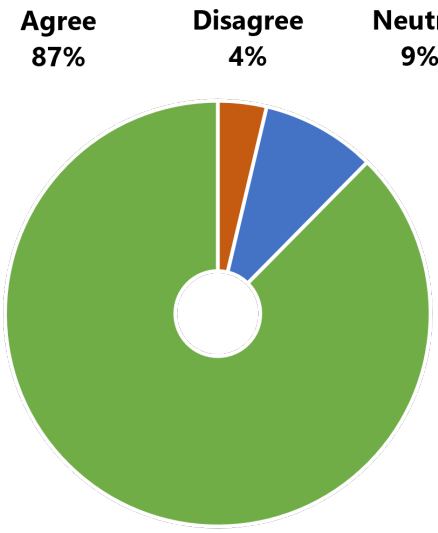
These observations suggest that large parts of the public find considerable value in the use of AI by police forces if it helps to protect vulnerable groups and society in a meaningful way.

'As far as catching a criminal quickly or finding a missing person is concerned, I think that any sensible person would vote in favour of using AI.'
– Estonia

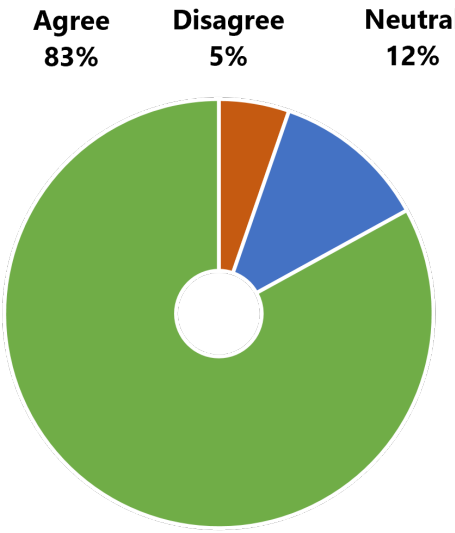
THE SECURITY OF SOCIETY PROFITS GREATLY FROM POLICE USE OF AI



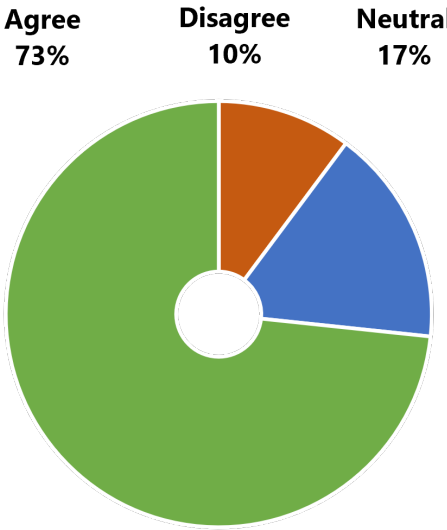
POLICE SHOULD USE AI TO SAFEGUARD CHILDREN AND VULNERABLE GROUPS FROM EXPLOITATION



POLICE SHOULD USE AI TO DETECT CRIMINALS AND CRIMINAL ORGANISATIONS

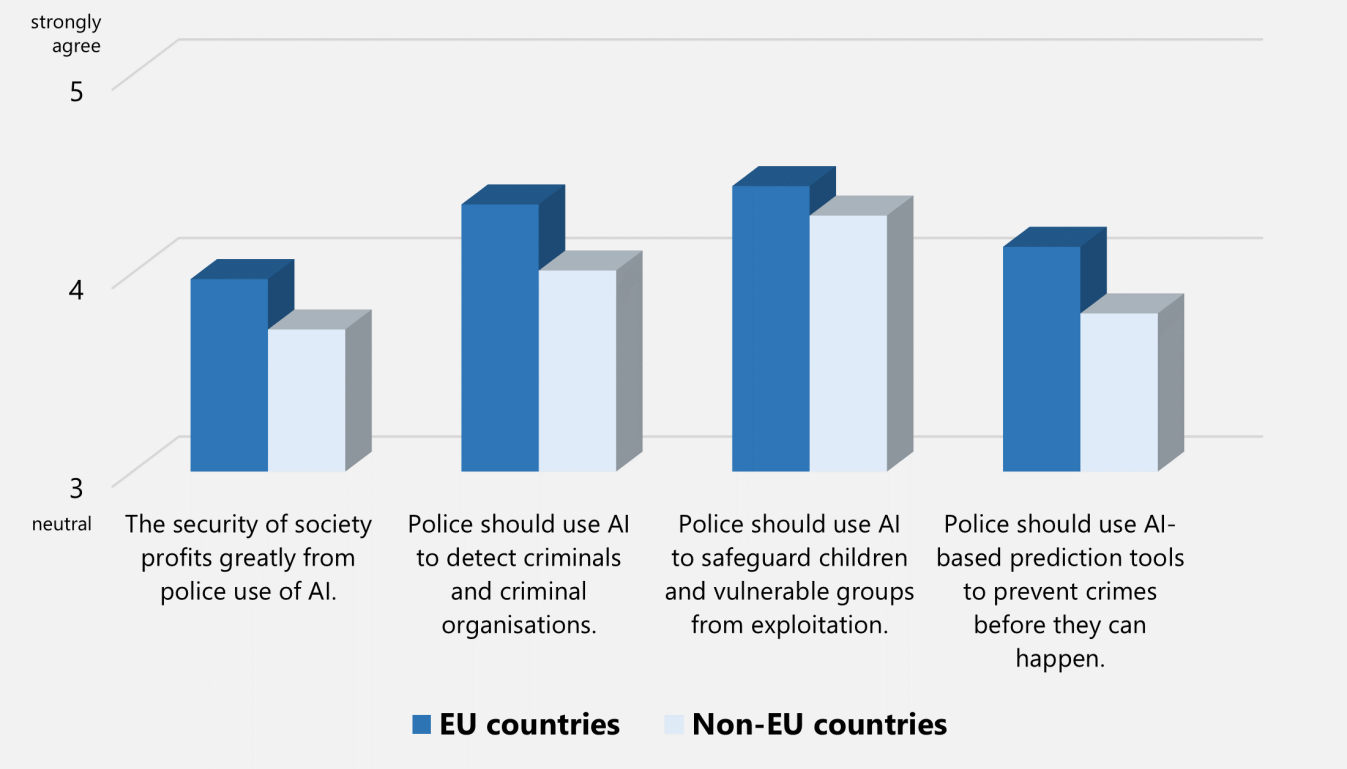


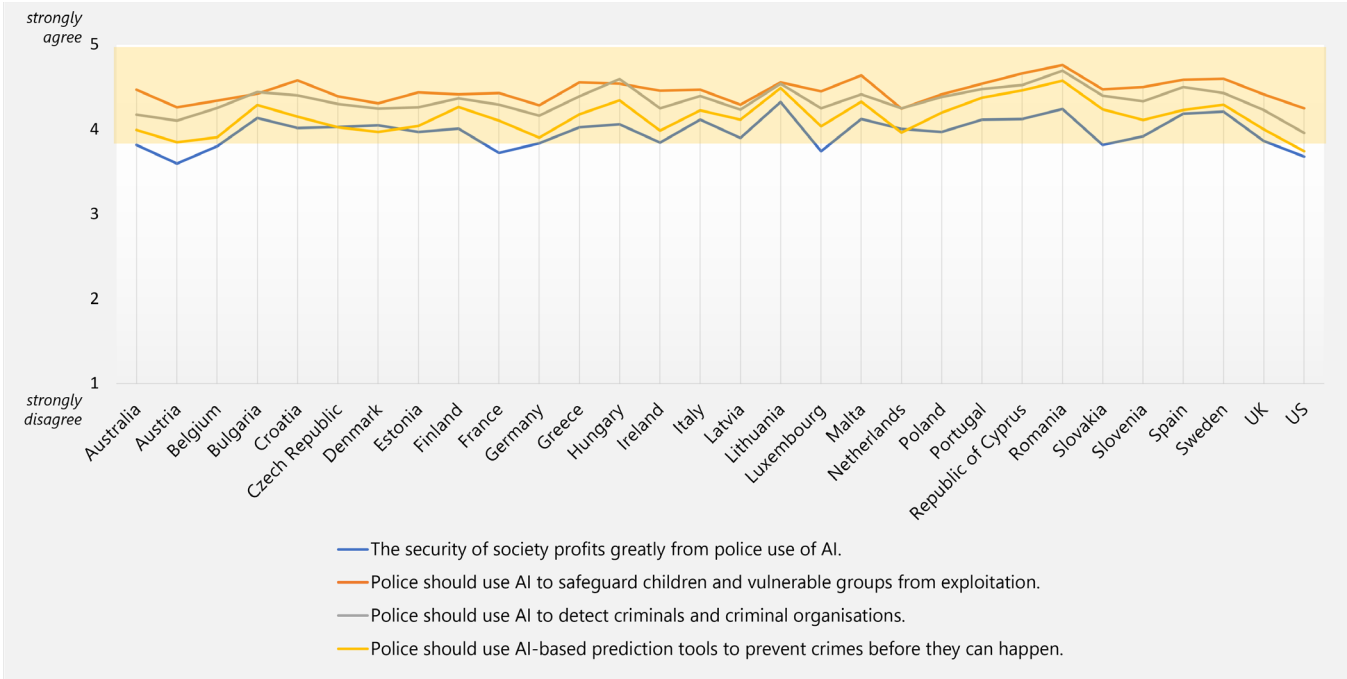
POLICE SHOULD USE AI-BASED PREDICTION TOOLS TO PREVENT CRIMES BEFORE THEY CAN HAPPEN



Country differences

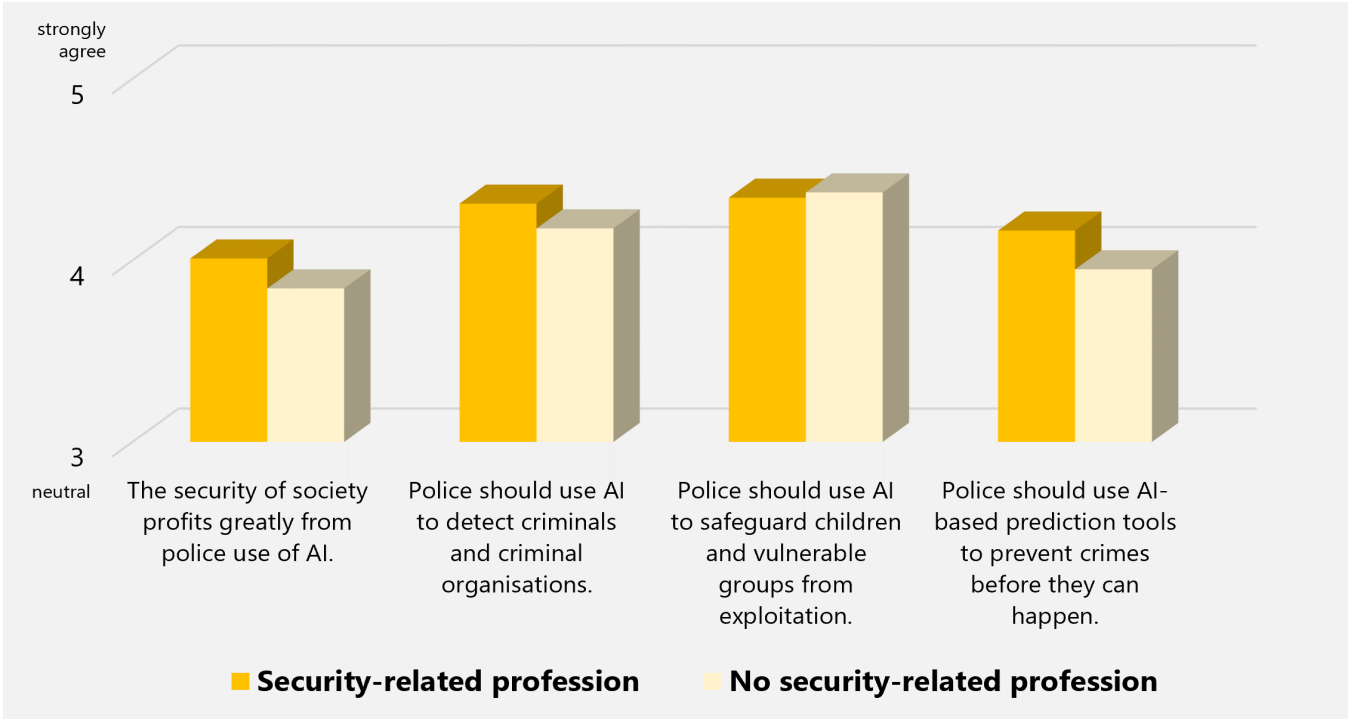
Participants in EU Member States demonstrated a slightly more positive perspective towards AI use by police compared to non-EU countries (UK, US, Australia). However, on average opinions were positive in all countries, with citizen reactions across countries presenting few variations.

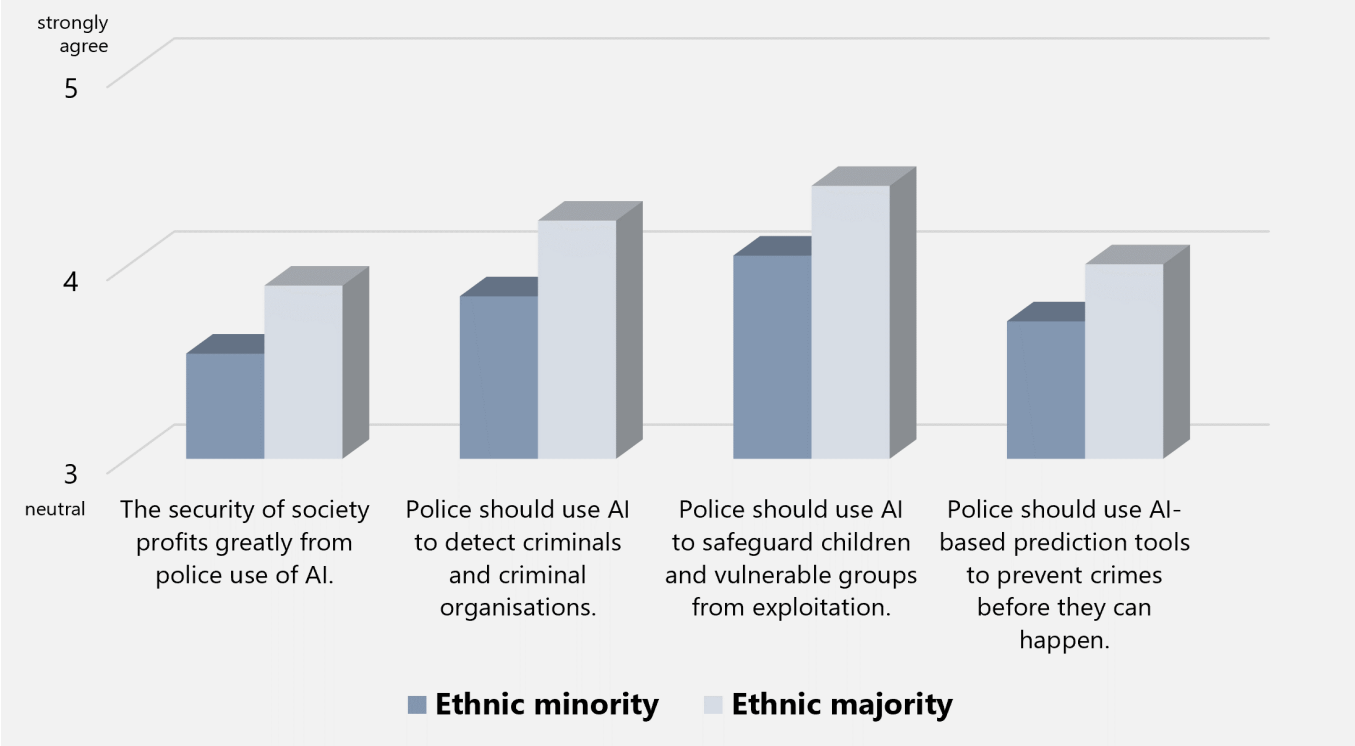




Group differences

Compared to men, women expressed a stronger wish for AI deployments if AI was used to safeguard children and vulnerable groups or to predict crime. Most positive towards all areas of AI use by police were people aged 55+, with younger cohorts slightly more critical. Working in a security-related profession was related to a stronger wish for AI use by the police, as was expertise in AI and expertise about AI use by police. In contrast, participants self-identifying as ethnic minorities were more critical of AI benefits than people who did not.





Additional participant observations

'I am not a specialist in this field, but I believe that the development of AI could improve policing in a real and direct way, so that everyone has the confidence that policing really guarantees security.' - Latvia

In their free-text comments, participants confirmed their overall support for AI use by police forces, highlighting its key benefits around the prevention and early detection of crime and security more generally. AI is described as a '*worthwhile tool*' that can help police have a positive impact on society, with some stating that '*it would be a good advantage against some criminal organisations*' and that *'it could be used to prevent gang-related or drug-related crimes'*. Generally speaking, respondents tended to acknowledge that there is some AI use that has become '*necessary*' and that '*not using AI is a risk*', particularly because it would place police at a disadvantage.

Answers show that, when participants think about the positive implementation of AI by police forces, they picture the protection of vulnerable citizens such as children, protection against crime, cybercrime, scammers, and corruption, or even help with finding missing people. However, most comments about the benefits of AI frame these within the limits of a particular purpose and indicate that AI can be beneficial, although this is not a given. Participants use phrasings such as '*if used correctly*' or '*AI can be...*', displaying an awareness of potential misuses and the need for rules.

CONCERNS ABOUT NEGATIVE CONSEQUENCES AND PRIVACY

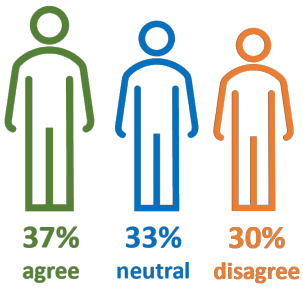
While our findings indicate an overall positive stance about the benefits of AI use by police forces, they also reflect concerns about privacy and potential negative consequences. Worries about possible **negative effects due to police decisions based on AI** were noted by 37.2% of respondents, with the remaining participants indicating either a neutral stance (33.0%) or no fear about potential negative effects (29.8%).

Nearly half of the participants voiced **concerns about police using AI to monitor their behaviour**, which were slightly higher for **online information** compared to **offline activities** (48.7% versus 46.7%). This suggests that about half of the participants had at least some worries about their privacy online or offline. About a third of participants remained neutral, while about 20% disagreed or strongly disagreed with being concerned. Relevant to note in this context is that the degree of privacy concerns shows a significant link to participants' perceived benefits of AI use by police and whether they expect police to deploy AI for the three specific security areas in our survey: the lower the level of privacy concerns, the more benefits ($r = -.26$, $p<.01$) and expectations to deploy AI ($r = -.28$, $p<.01$) we found.

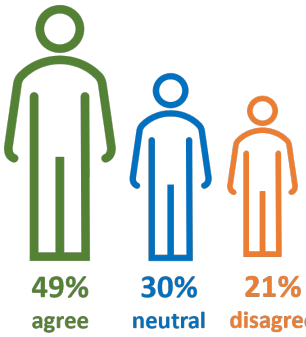
In addition, the more police forces were perceived to make **sufficient efforts to avoid negative consequences** and to **show respect** in their activities, the higher were the positive perceptions of AI and the lower were concerns about privacy and negative consequences. Credible efforts by police forces thus seem to provide reassurance. These observations speak for a clear **need to have adequate mechanisms** in place that give citizens trust about police forces' correct AI use.

In this respect, it is somewhat concerning that only half of all participants found that **police in their country make sufficient efforts** to avoid the negative consequences of AI. Further, only 60% of participants found that **police respect citizens' rights** in their activities.

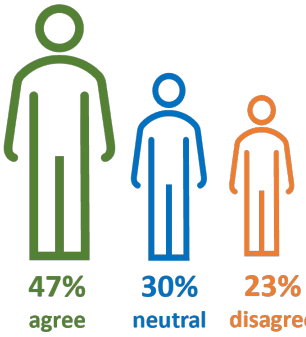
I AM AFRAID THAT POLICE DECISIONS BASED ON AI COULD WORK IN MY DISADVANTAGE



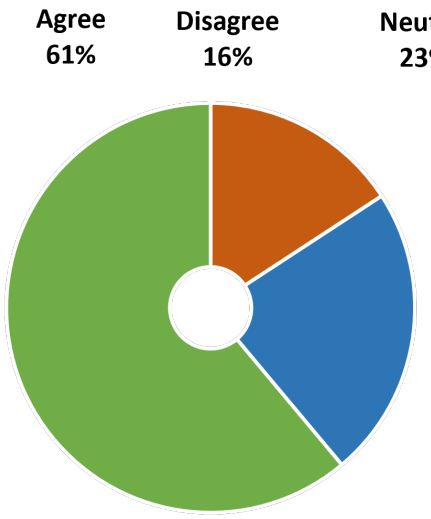
I AM CONCERNED THAT AI MAKES MY ONLINE INFORMATION MORE OPEN TO POLICE SCRUTINY



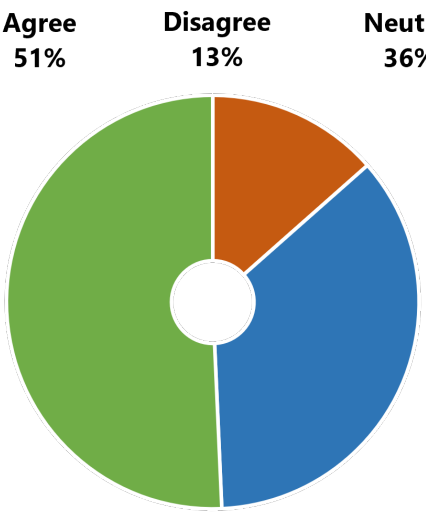
I AM CONCERNED THAT AI INCREASES POLICE POWERS TO MONITOR MY ACTIVITIES OFFLINE



IN THE EXECUTION OF THEIR TASKS, POLICE RESPECT CITIZENS' RIGHTS

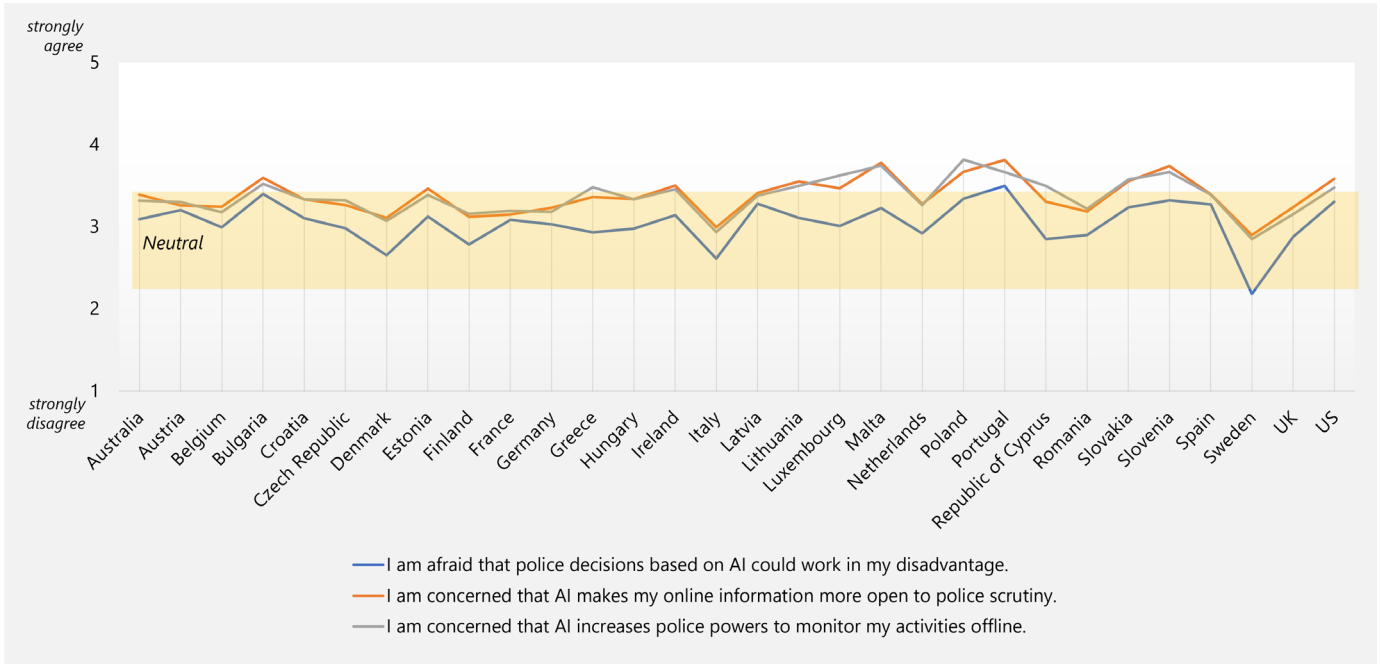


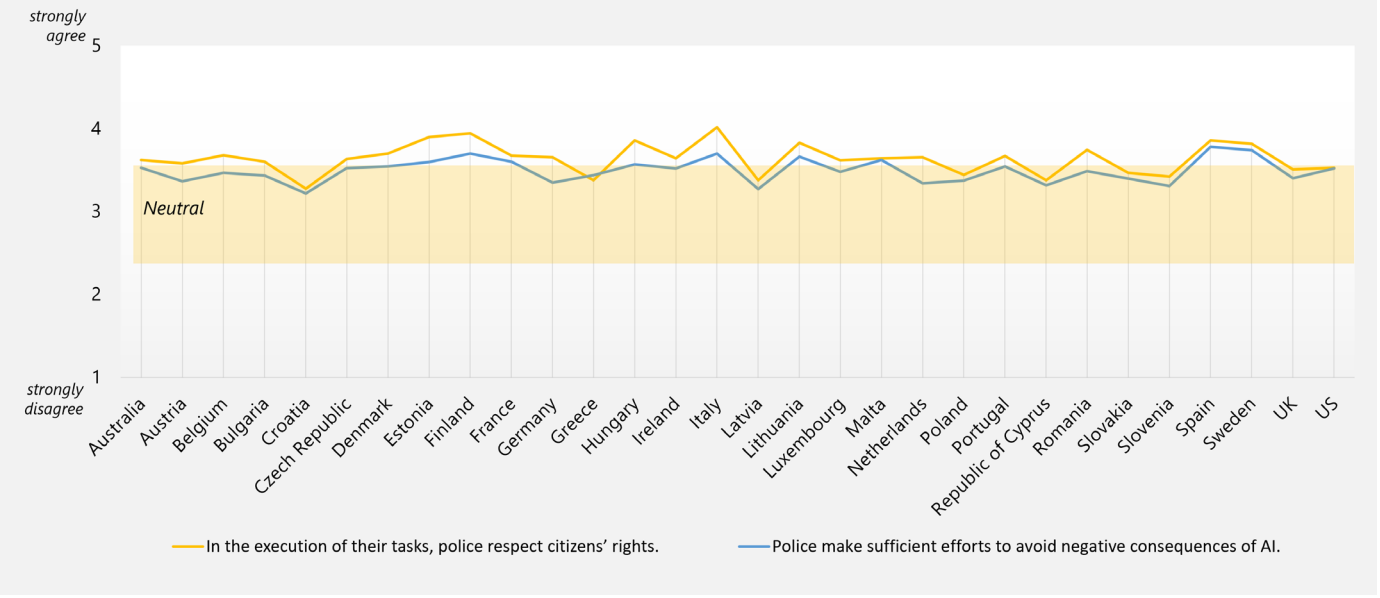
POLICE MAKE SUFFICIENT EFFORTS TO AVOID NEGATIVE CONSEQUENCES OF AI

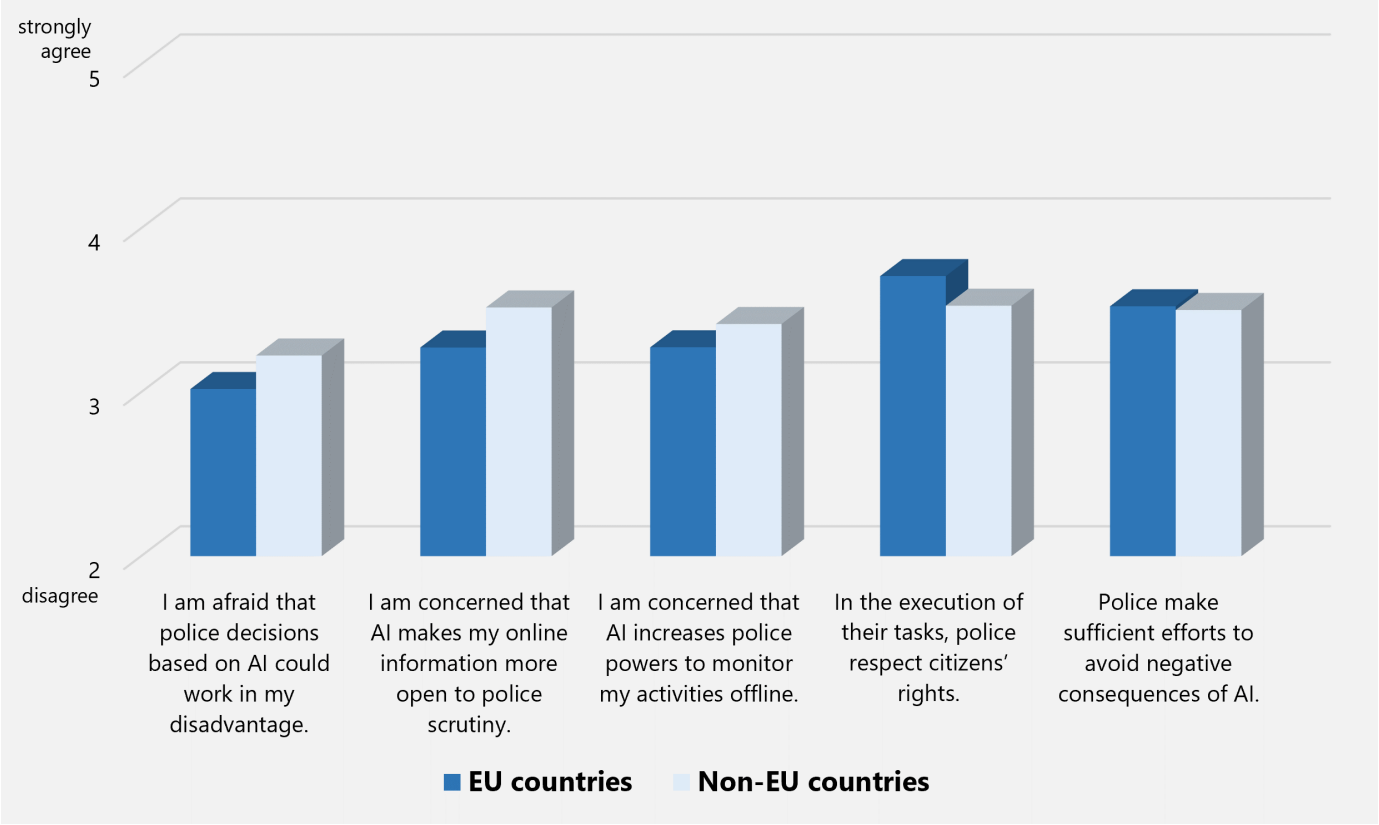


Country differences

The level of concern across countries was on average moderate and differed only marginally. Comparing EU with non-EU countries shows a slight tendency for citizens in EU countries to report a lower level of concerns and somewhat more positive perceptions of police in their efforts to respect citizens' rights and avoid negative consequences of AI, but again, differences remain small.





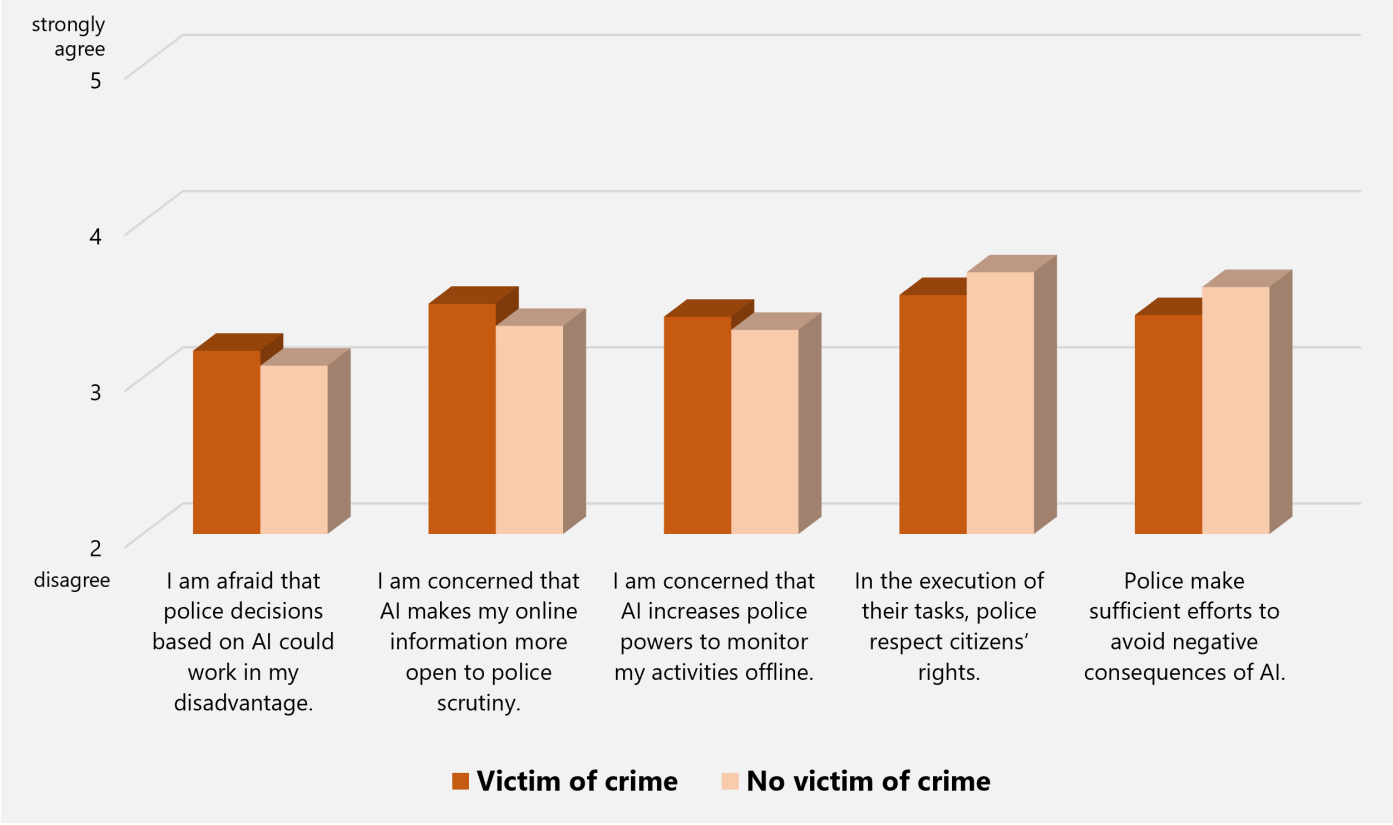


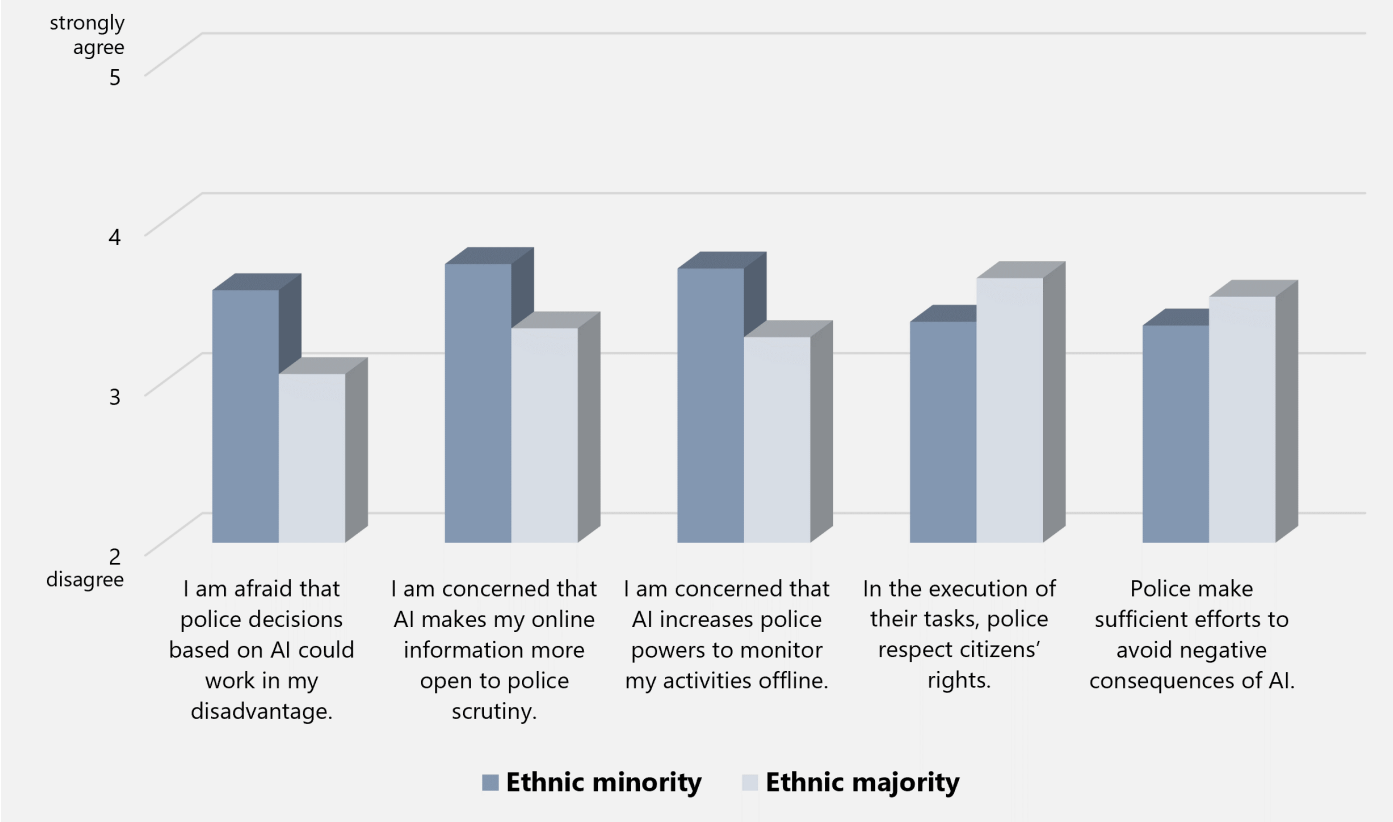
Group differences

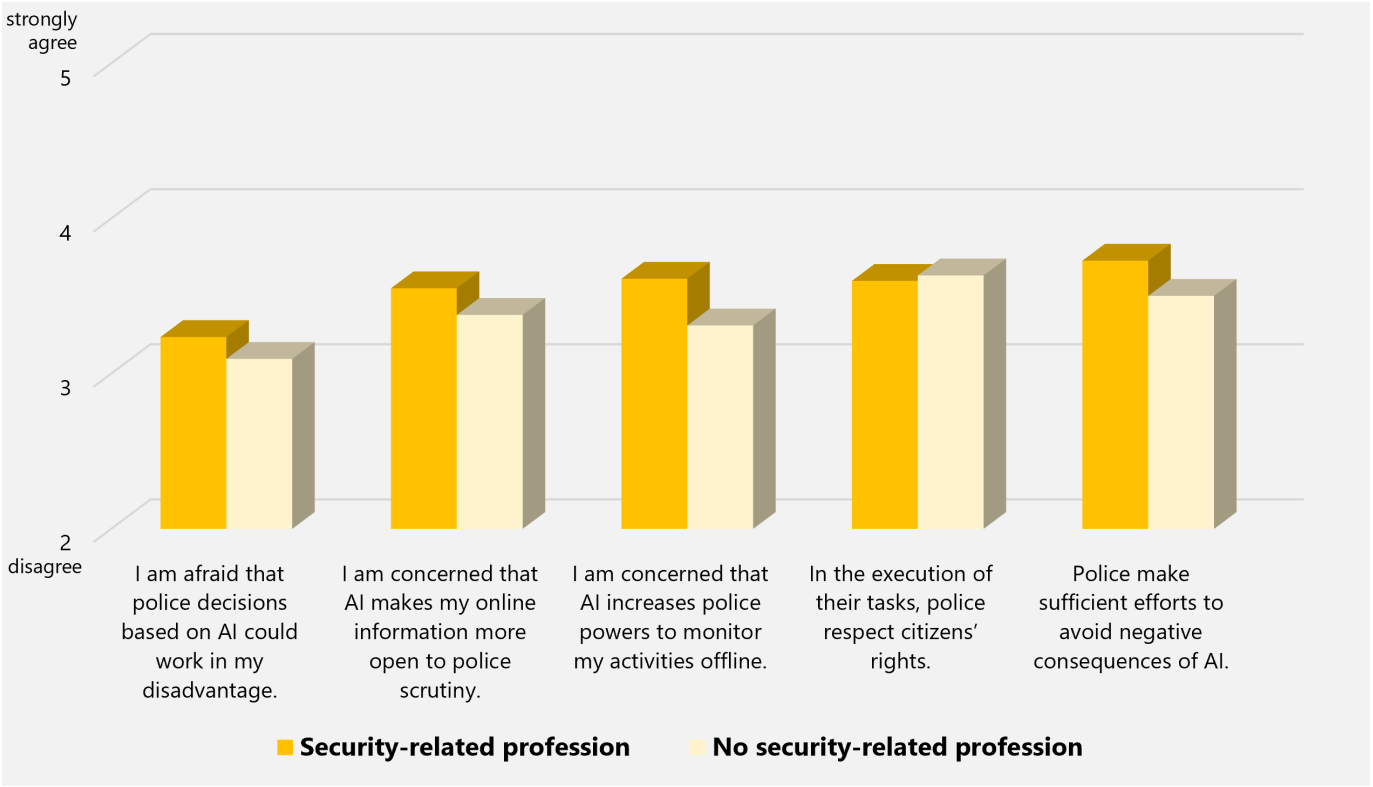
Women voiced slightly fewer concerns and more positive views about respect for and handling of AI consequences compared to men. Age groups showed a tendency to decrease concerns with increasing age, i.e., y**ounger age groups were slightly more concerned** than older participants and at the same time voiced a more negative outlook towards police. This is in line with the stronger perception of AI benefits found by participants 55 years of age or older.

Participants **self-identifying as ethnic minorities** were again more critical towards AI, with higher concerns and less positive perspectives towards police. Higher concerns were also reported by individuals with past **victimisation experience from crime**.

In contrast, participants in a **security-related profession voiced stronger concerns** but at the same time also showed a stronger conviction that police forces do enough to avoid the negative consequences of AI.


strongly agree
5
4
3
2
disagree
I am afraid that police decisions based on AI could work in my disadvantage.
I am concerned that AI makes my online information more open to police scrutiny.
I am concerned that AI increases police powers to monitor my activities offline.
In the execution of their tasks, police respect citizens' rights.
Police make sufficient efforts to avoid negative consequences of AI.
Victim of crime
No victim of crime


strongly agree
5
4
3
2
disagree
I am afraid that police decisions based on AI could work in my disadvantage.
I am concerned that AI makes my online information more open to police scrutiny.
I am concerned that AI increases police powers to monitor my activities offline.
In the execution of their tasks, police respect citizens' rights.
Police make sufficient efforts to avoid negative consequences of AI.
Ethnic minority
Ethnic majority



Additional participant observations

Comments provide further insights into the concerns participants have with regard to police use of AI. **Privacy** is one of the key worries, with participants wondering what kind of data is gathered and how it could be used and misused. By way of illustration, one respondent compared AI use with 'Big Brother' watching over citizens. There are strong concerns in the process of handling potentially sensitive data about the **protection of the rights of those who are being investigated**. Respondents also raised questions around the length of **data storage, confidentiality, data access requirements**, and the wider ethics around producing AI data.

'The problem is a permanent surveillance of an ordinary citizen, in order to avoid crimes.'
– Belgium

Another key concern respondents expressed is the **potential misuse of AI**, particularly through negative influences over its use and the possibility of biased outcomes. Concerns were also raised about **potential biases** that could lead to discriminatory practices. Respondents listed the profiling of ethnic minorities and foreign nationals as examples.

'It is important that interested groups are not able to influence police use of AI.'
– Lithuania

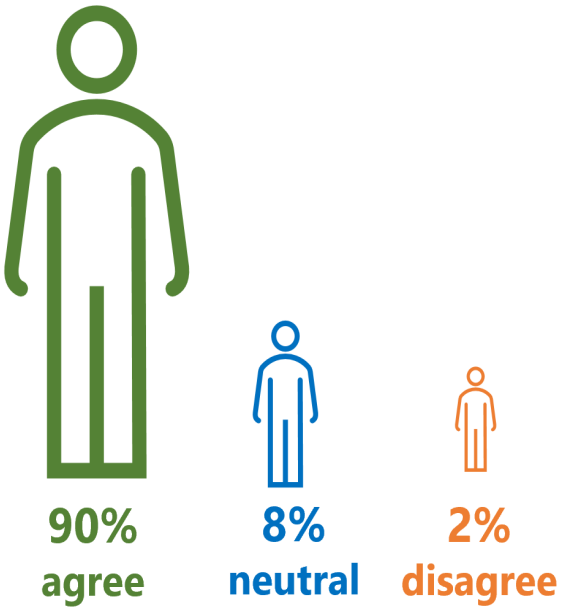
A number of respondents further insisted that AI use **must not fall into the wrong hands**, where it could be liable to exploitation by people with political power for their own projects or be deployed under the influence of particular governments. Oftentimes, the influence of politics was viewed in the realm of '*corruption*', with respondents worrying about corruption among party politics leading to potential bias and misuse of AI at the hands of police forces. Likewise, various respondents expressed mistrust in their regional or national police forces, which was accompanied by a concern for potential AI misuse. The additional observations thus emphasise again the relevance of trust in the police.

CLEAR NEED FOR BETTER AI ACCOUNTABILITY

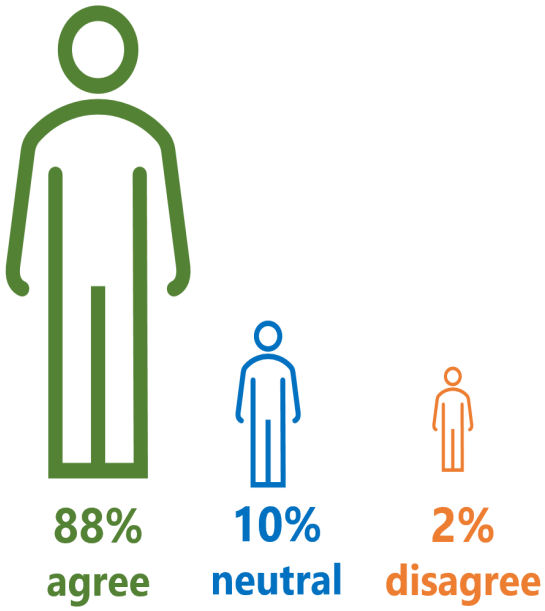
Asked about a need for AI accountability, participants felt strongly that police should be held accountable for their AI use: 90% expect police to be **held accountable for the way they use AI**, and 88.3% **for the consequences of their AI use**. This suggests that citizens expect strong mechanisms as well as reassurance that the police are willing to deploy AI in an appropriate way.

The suggestion of an **overarching AI Accountability Framework found broad citizen approval**: 81.8% found a universal Accountability Framework important or very important (compared to 15.1% neutral and 3.1% low or no importance).

POLICE SHOULD BE HELD FULLY ACCOUNTABLE FOR THE MANNER IN WHICH THEY USE AI

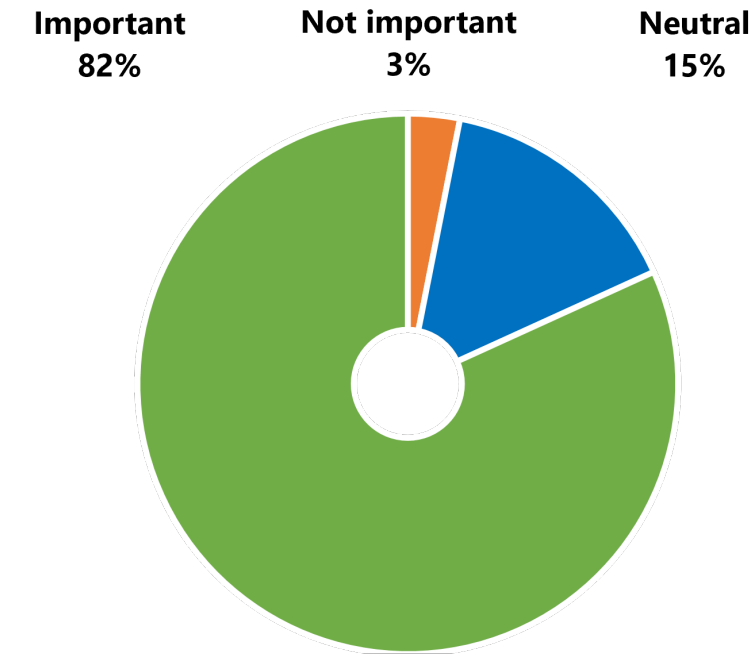


POLICE SHOULD BE HELD FULLY ACCOUNTABLE FOR THE CONSEQUENCES OF THEIR AI USE

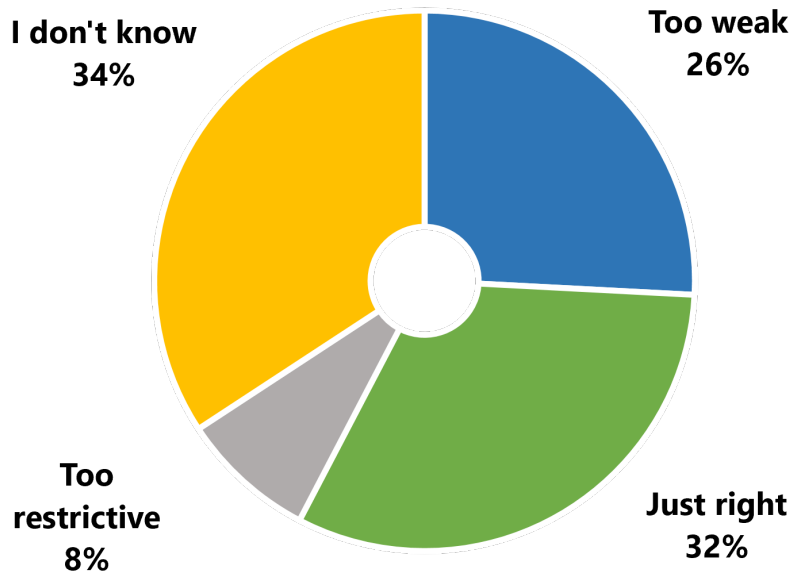


However, **only a third of participants considered existing mechanisms appropriate**. 25.8% perceived them as too weak, while 8.1% rated them as too restrictive. At the same time, we found that a considerable number of participants reported that they "don't know" whether current accountability mechanisms are appropriate (34.2%). This implies that a substantial part of the public lacks sufficient information about existing mechanisms to make an informed judgement—a clear call to action to improve information and transparency.

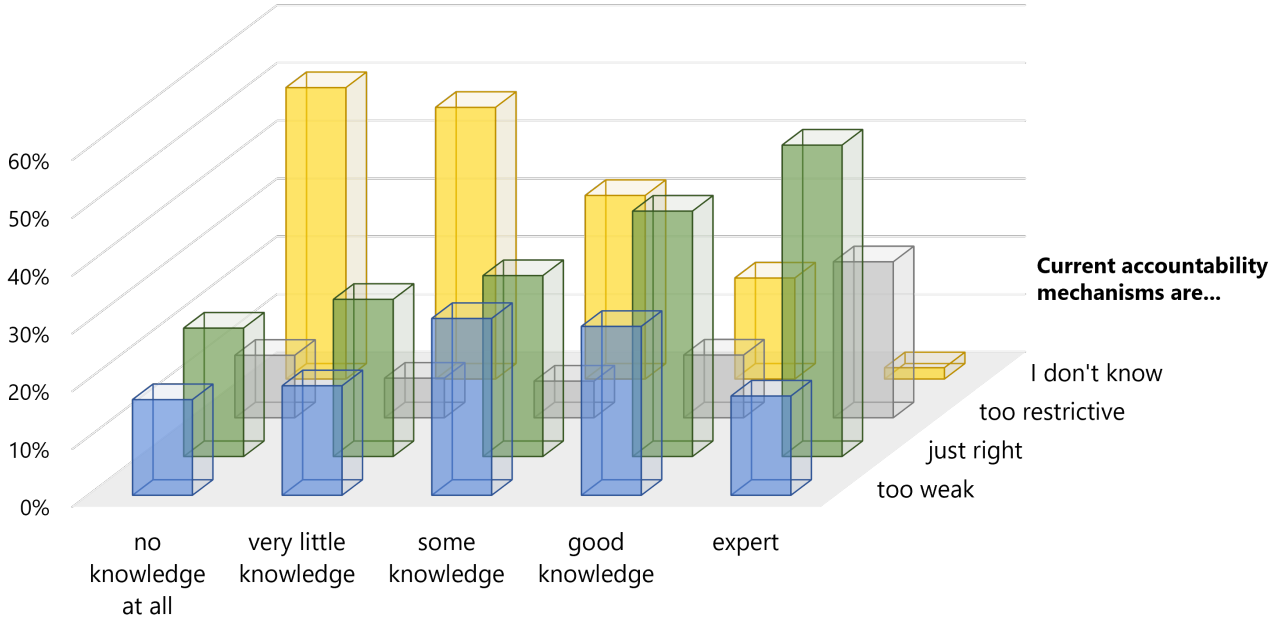
HOW IMPORTANT IS IT THAT A UNIVERSAL FRAMEWORK IS CREATED THAT ENSURES THE ACCOUNTABILITY OF AI USE BY POLICE?



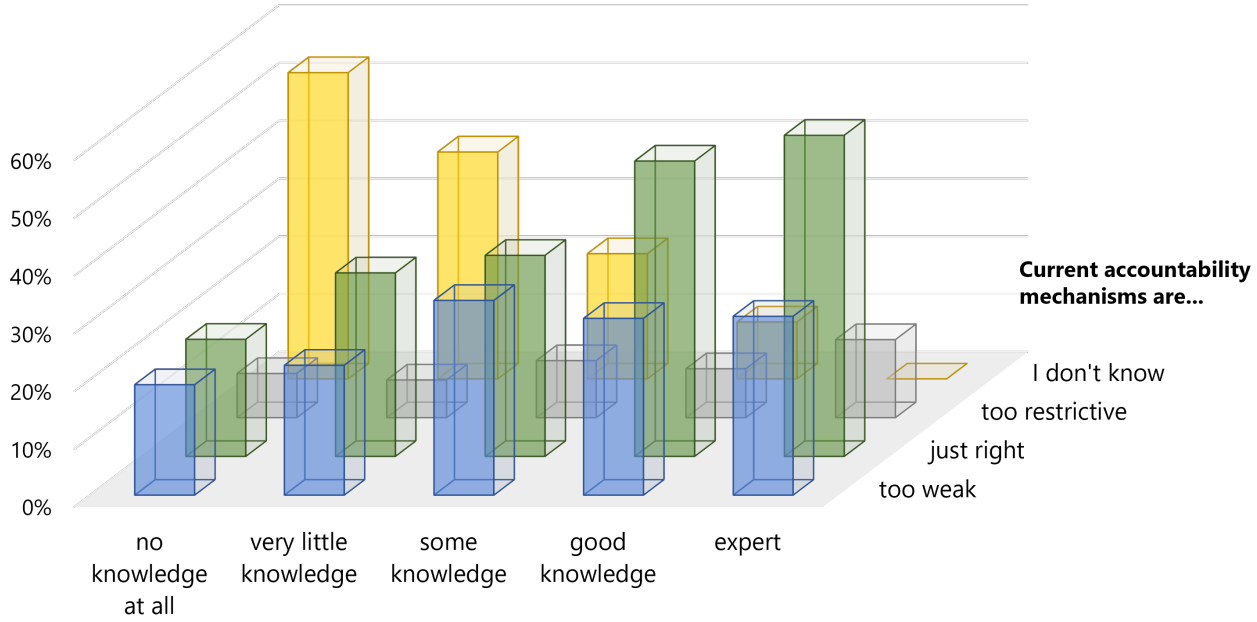
THE CURRENT MECHANISMS FOR HOLDING POLICE ACCOUNTABLE FOR THEIR AI USE ARE...



HOW MUCH DO YOU KNOW ABOUT AI IN GENERAL?



HOW MUCH DO YOU KNOW ABOUT THE USE OF AI BY POLICE?



A **comparison of participants with different levels of AI knowledge** reveals interesting differences between experts and people who claim to have little or no knowledge. The percentage of participants rating accountability mechanisms as 'too weak' is similar across all knowledge levels. However, participants who perceived themselves to have little or no knowledge about AI overwhelmingly claimed to 'don't know whether accountability mechanisms are appropriate (48.0%). In contrast, people who claim good AI knowledge or even AI expertise were very likely to perceive current mechanisms as 'just right' (43.3%). The same picture arises when comparing disparate levels of knowledge about AI use by police, although the pattern is here even more extreme.

Participants with high general AI knowledge were also more likely to rate existing mechanisms as 'too restrictive' (12.2%) when compared to other knowledge levels (some knowledge: 6.3%; no/little knowledge: 8.0%). Interestingly, participants with expertise in police-related AI use did not show the same preference with 9.3% rating current mechanisms as too restrictive, compared to some knowledge: 9.9%, no/little knowledge: 7.0%. However, experts in police-related AI use more often experienced existing mechanisms as too weak.

Additional participant observations

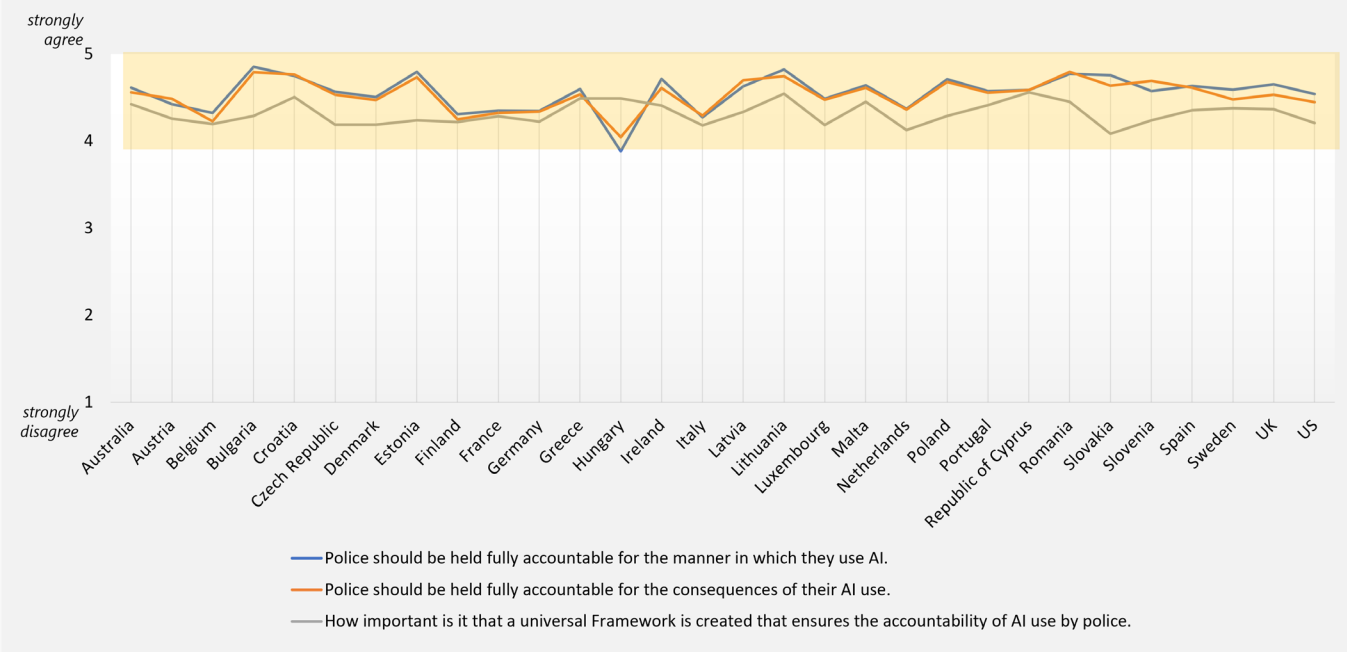
Reviewing participants' comments added some important nuances, specifically a strong need for transparency about AI deployments. For example, participants insisted that police '*must prove that [AI] works well and is risk-free*' and that its use is '*open and verifiable*'.

'AI is a good tool to tackle serious crime. However, accountability must be provided to society about its use, and this must be accessible.' – Netherlands

Participants repeatedly used words such as '*open*', '*clear*', '*impartial*', '*verifiable*', '*accountable*', and '*legitimate*' to describe a positive and transparent vision of AI use by police forces. Furthermore, respondents showed a desire to go beyond being simply '*informed subjects*' of AI use and instead called for continuous citizen involvement in AI deployment. This links participants' expectations to the AP4AI Principles of Explainability, Conduct, Transparency, and Pluralism to ensure that the public is aware of its capabilities and implementation, that conduct is adequate or otherwise disciplined to minimise wrongdoing and that stakeholders are involved from the first stages.

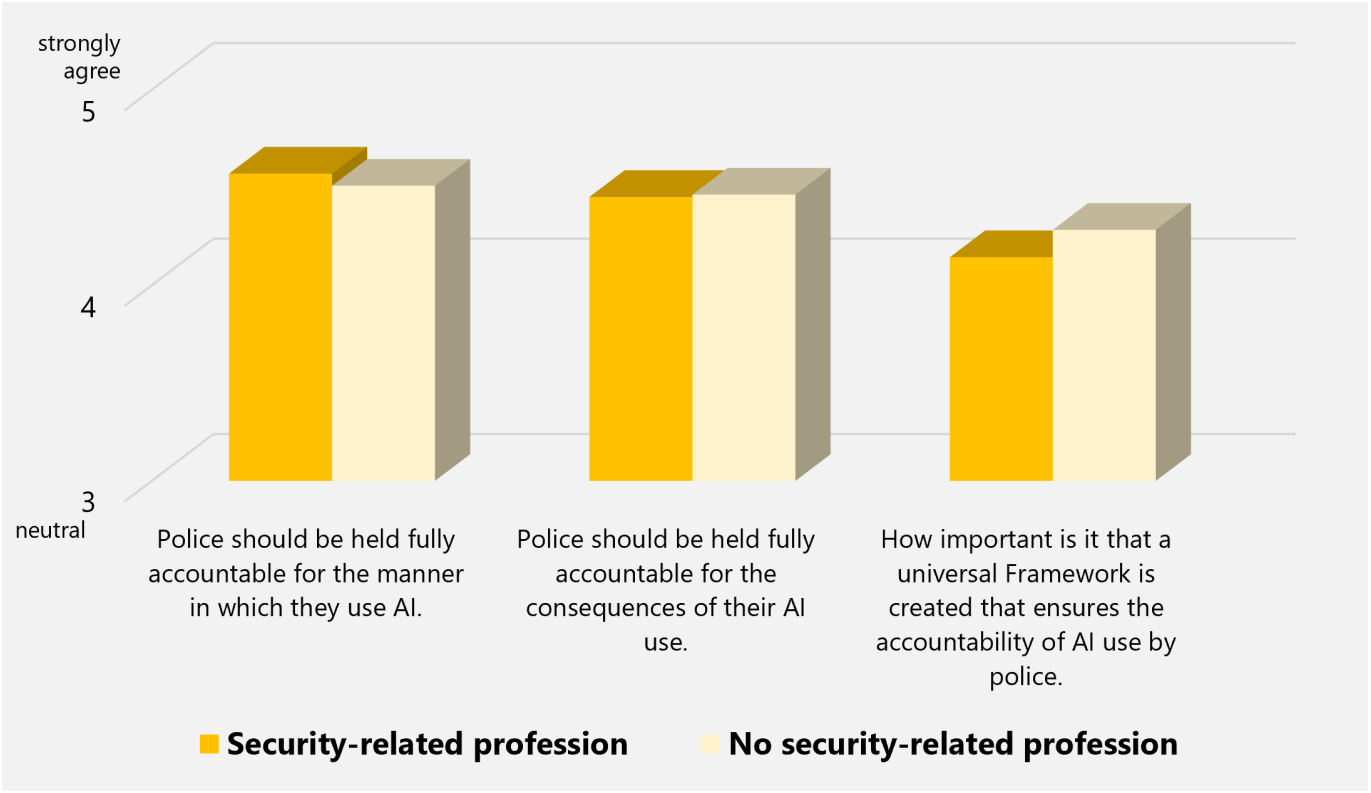
Country differences

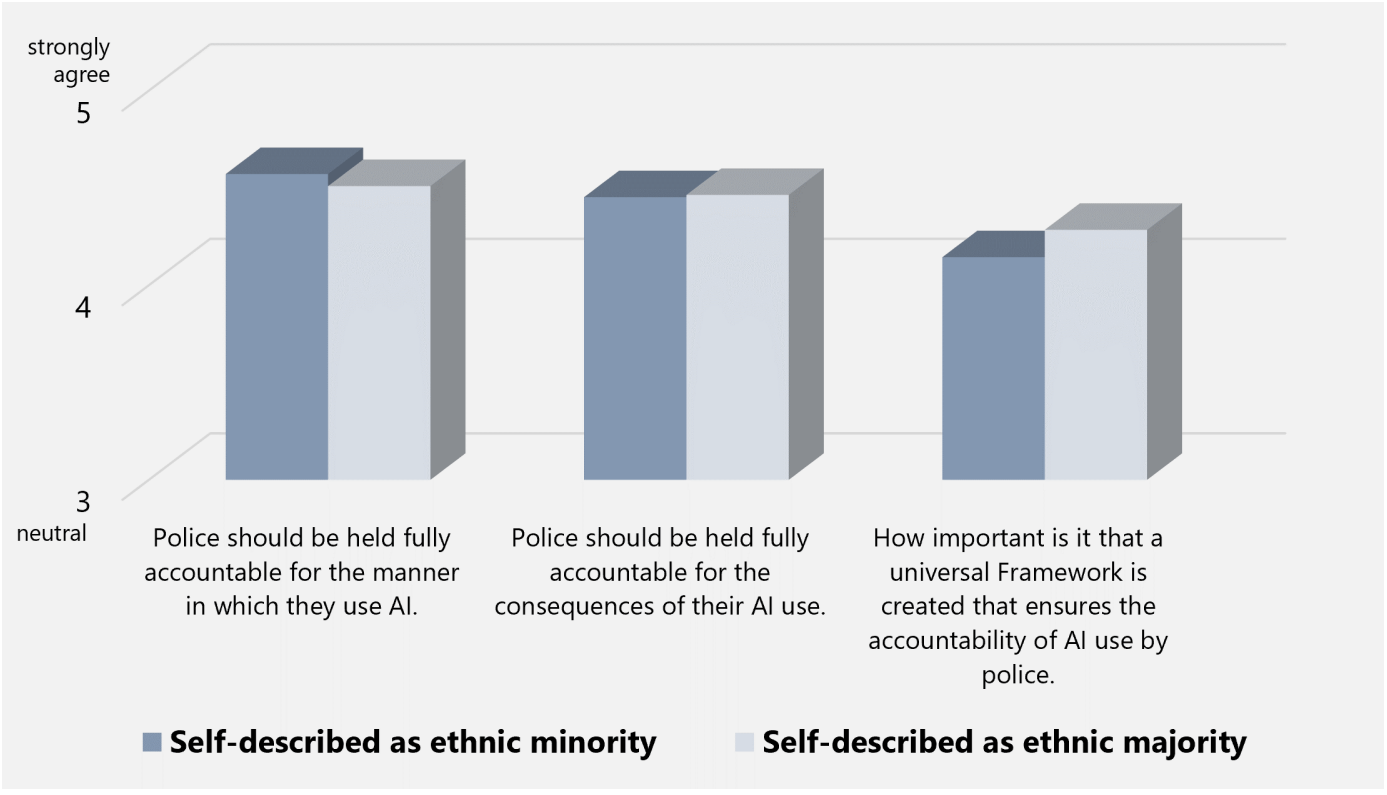
In all 30 countries, participants not only agreed that police should be held responsible but also that a universal framework would be important. In this case, participants from EU and non-EU countries did not differ, representing a comparable degree of appetite for strong accountability mechanisms.



Group differences

All genders and age groups expected a high degree of AI accountability, although men indicated somewhat higher expectations than women, and expectations tended to increase across age groups, with the highest level for people 55 years and older. The differences, however, were small. The same is true for differences between professional groups, members of ethnic minority and majority groups, and people with or without crime victimisation experiences. These observations illustrate that AI Accountability is of high importance across all demographic groups.


strongly agree
5
4
3
neutral
Police should be held fully accountable for the manner in which they use AI.
Police should be held fully accountable for the consequences of their AI use.
How important is it that a universal Framework is created that ensures the accountability of AI use by police.
Security-related profession
No security-related profession


strongly agree
5
4
3
neutral
Police should be held fully accountable for the manner in which they use AI.
Police should be held fully accountable for the consequences of their AI use.
How important is it that a universal Framework is created that ensures the accountability of AI use by police.
Self-described as ethnic minority
Self-described as ethnic majority

CITIZEN EXPECTATIONS ON HOW TO ENSURE AI ACCOUNTABILITY

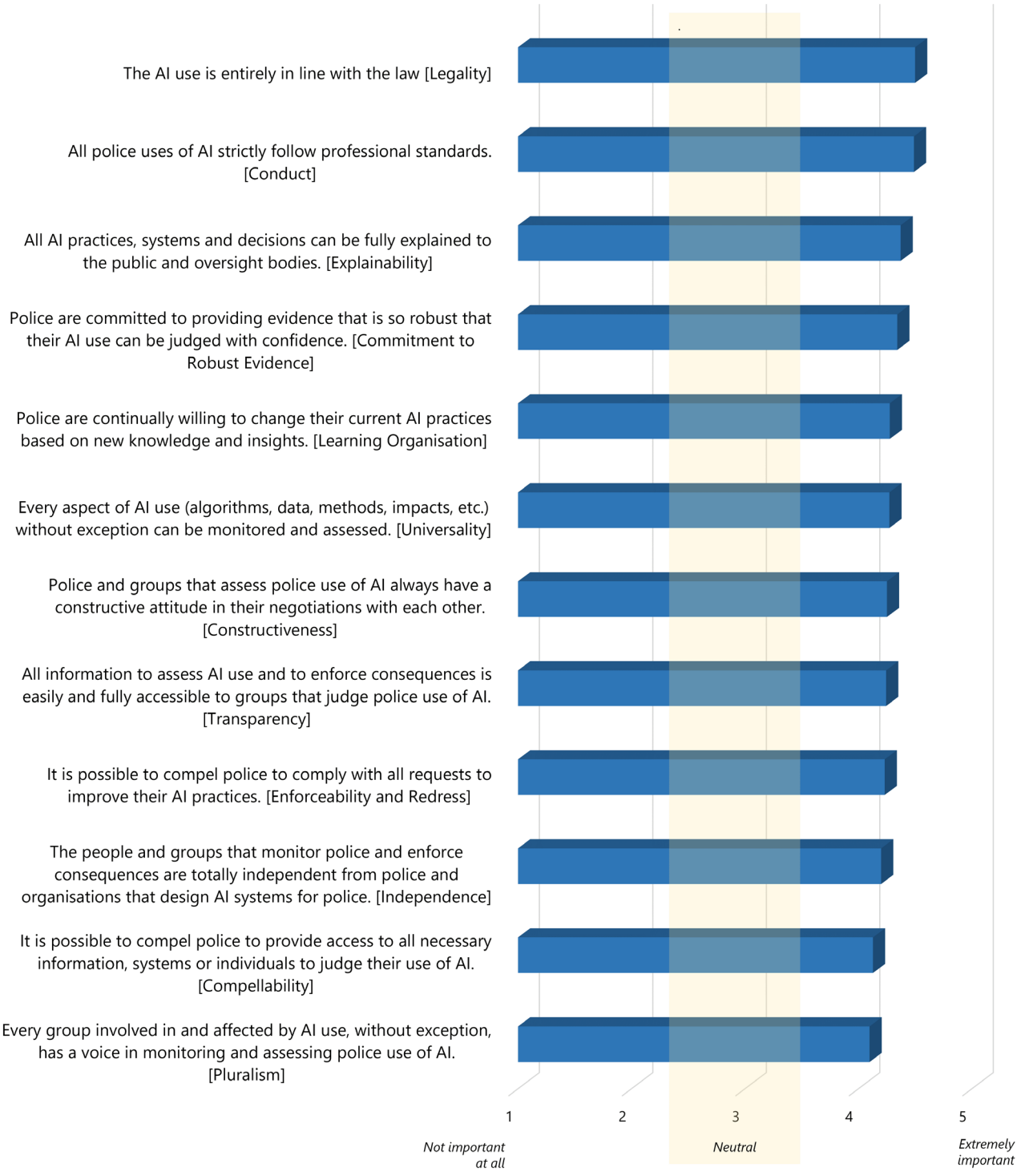
AP4AI puts forward 12 principles that together constitute AI Accountability. The citizen consultation allowed us to evaluate to what extent citizens consider these principles important to make them feel confident about the use of AI by police.

Relevance of AP4AI Principles as a mechanism to ensure AI Accountability

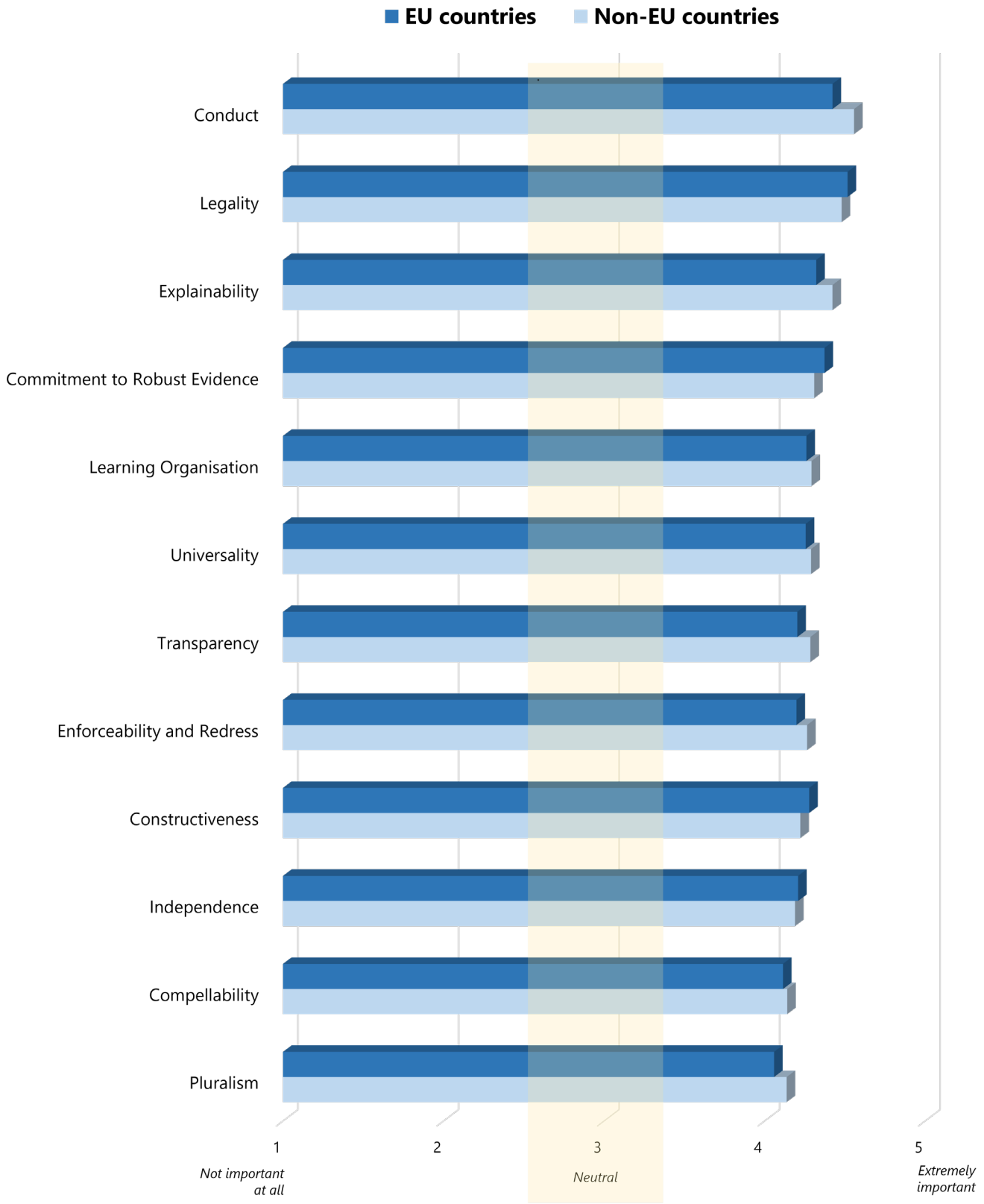
All 12 principles were validated as important for AI Accountability. While Legality, Conduct and Explainability emerged as the most important principles, we only found minor variations among the 12 principles. A more detailed view of the importance of principles as rated in EU countries versus non-EU countries shows some disparities in the order of the 12 principles. However, the observed differences between principles are small and therefore not necessarily meaningful in a practical sense. What the comparison can demonstrate is that the importance of specific principles may vary depending on factors such as communities and contexts in which AI capabilities are deployed (cf. also Akhgar et al., 2022b).

Overall, the results validate the overall relevance of all 12 AP4AI principles as a meaningful foundation for an AI Accountability Framework.

HOW IMPORTANT IS IT THAT...



HOW IMPORTANT IS...



Citizen proposals to implement the AP4AI Principles

The citizen consultation provided participants with the opportunity to suggest additional mechanisms and safeguards that will give them confidence that the police use AI in an appropriate way. Nearly 50% of participants used this opportunity to provide free-text input (a total of 2,552 entries).

The review of the answers revealed that recommendations fell into the remit of the AP4AI principles in one form or another, which validates the capacity of the 12 principles to provide an acceptable framework for citizens to ensure trust in AI deployments by the police.

Equally important, however, is to review the manner in which citizens reflect the 12 principles. Such details provide important information about the way citizens want the principles to be implemented and, as such, can inform the mechanism with which the AP4AI approach is transferred into operational practice.

The following sections thus look at how respondents' comments fit within and further detail each of the AP4AI principles.

Legality: *AI use is entirely in line with the law*

A legal framework was one of the key concerns for respondents, who noted that specific legislation must be developed to regulate AI use by police forces. They further stressed that compliance with these regulations must be ensured. This means adding to existing AI legislation while also developing new frameworks that address new challenges. Participants noted that lawful use of AI has the potential to avoid concerns and make them '*feel safe*'. A few participants specifically stated the need for legislation to precede AI deployment, with one addressing both regulations and consequences of misuse: '*There should be a framework around which the use of AI by the police is based, until it can be implemented, with sanctions and justice when it is abused*'.

'A strict legal framework and solid protection of all data in accordance with the GDPR and destruction when it does not apply within a criminal framework.'
– Netherlands

Selected answers referencing Legality	
Legality should be guaranteed by	
• Laws and regulations that are designed for AI implementation • Clear, simple, and strict laws • Laws that are equally applied to everyone • Implementation of laws and regulations before AI is deployed • Having clearly defined rules on the responsibilities and rights of police officers	• Laws and regulations that cover the rights of both police officers and citizens • Focusing on the protection of citizens • Focusing on Prevention of Misuse • Designating independent oversight (link to the Independence Principle)

'Everything could be open and not so secretive.'
– Estonia

Universality: *every aspect of AI use (algorithms, data, methods, impacts, etc.) without exception can be monitored and assessed*

Compared to other principles, aspects of Universality were less frequently commented on. Where Universality was addressed, respondents primarily showed a desire for the ability to monitor and assess all aspects of AI use.

Selected answers referencing Universality	
Universality should be guaranteed by	
• Being open and not secretive • Being fully accountable and transparent	• Ensuring complete oversight • Allowing access to all data and procedures

'The people's right to know and to participate in the functioning of AI and the right to participate in the definition of the rules governing AI.'
– Slovenia

Pluralism: *every group involved in and affected by AI use, without exception, has a voice in monitoring and assessing police use of AI*

Pluralism was reflected primarily in terms of the disparate communities that might be negatively affected by security-related AI deployments, often with a focus on vulnerable or disenfranchised groups. This also found expression in calls for '*equality*' and the avoidance of bias and misuse of AI within the development and deployment of AI tools by police. Concretely, participants called for more thorough ways to address the fair use of AI that do not allow for specific social groups to be worse off by its use. This means acknowledging the effect biases can have on the design and applications of AI. Participants thus emphasised Pluralism primarily in terms of assurance that AI use does not discriminate, particularly against minorities.

Selected answers referencing Pluralism	
Pluralism should be guaranteed by	
• Not targeting minorities disproportionately • Making sure use is fair to everyone	• Treating everyone equally • Providing access to AI use and data to lawyers, courts, and beyond

'Transparency in their actions and, above all, in the limits to which they can go; always bearing in mind as citizens that the information relating to their actions must be confidential for greater public safety.'
– Lithuania

Transparency: *all information needed to assess AI use and enforce consequences is easily and fully accessible to groups that judge police use of AI*

Participants repeatedly expressed their desire for the transparent use of AI through the provision of clear information regarding its deployment by police: '*All steps should be disclosed transparently*'. This, should also include 'how and when AI is used'. Participants thus expressed an urgent need for police to communicate transparently what they do and why.

Selected answers referencing Transparency	
Transparency should be established by	
• Honest and full disclosure of use, capabilities, and limitations, including when, how, and why it is used • Disclosure of people operating AI tools • Documentation of all uses • Availability for auditing purposes • On request, data about a person should be made available	• Disclosure: how data is handled • Disclosure of whether use is truly beneficial • Police transparency in cases of wrongdoing • The publication of irregularities • Continuous monitoring

***Independence**: the people and groups that monitor police and enforce consequences are totally independent from police and organisations that design AI systems for police*

Independence was a recurrent topic with a focus on the independence of monitoring, control, and oversight. Respondents called for oversight bodies that include human supervision of AI and independent bodies to monitor AI usage by police. One participant captured this expectation by requesting '*that the people and groups that monitor police and enforce consequences are totally independent from police and organisations that design AI systems for police*'.

'Whatever legislation and investigative powers are set, they should be monitored and enforced through organisations that are independent of the police.'
– UK

Participants also offered recommendations for where responsibility should lie. The list that emerges is large and varied, with experts and bodies at various levels, including commissioners, lawyers, ministries, prosecutors, judges, civilians, AI specialists and scientists, dedicated government departments, internal affairs, members of parliament, and psychologists. The independence of these experts or bodies was seen as key for accountability to be successful, with respondents emphasising the '*incorruptibility*' and political neutrality oversight bodies must exhibit. At the same time, respondents also listed groups that should be excluded from monitoring and assessing police use of AI, including private industry, or more specifically, those profiting from AI, governments, convicted individuals, extremist groups, government and political groups or parties, police forces or security departments, and religious groups (see also section on Parties Responsible for AI Accountability).

Selected answers referencing Independence	
Independence should be guaranteed through	
• Neutral and independent oversight • Independent oversight, verifying any evidence gathered using AI tools • The setting of ethical standards by an independent body • A licencing system for AI use	• Responsible experts and bodies at various levels, including commissioners, lawyers, ministries, prosecutors, judges, citizens, AI specialists and scientists, dedicated government departments, internal affairs, MPs, and psychologists

'A monthly, public report detailing how and why AI was used and if the outcomes were positive or negative.'
– Australia

Commitment to Robust Evidence: *the police are committed to providing evidence that is so robust that their AI use can be judged with confidence*

The Commitment to Robust Evidence is closely linked to other principles such as Explainability and Independence, which both ensure the possibility of reliable evidence being created and maintained. Participants addressed robust evidence mostly by requesting that police forces produce strong evidence to support any actions taken and that such evidence can then be scrutinised by an independent oversight body.

At the same time, the Commitment to Robust Evidence links in with respondents' calls for additional security protocols, including additional restrictions, highly trained staff, and thorough vetting procedures.

Selected answers referencing Commitment to Robust Evidence	
Steps to guarantee Robust Evidence	
• Police must produce robust evidence to support AI use; this should include evidence of successful past use • Ensure that evidence is real, concrete, and gathered before any decisions are made • Robust evidence should be gathered before AI tool implementation	• Evidence should not only be gathered via AI tools • AI data gathered should not be used as evidence for an offence • Ensure AI use is based on evidence to avoid profiling minorities • Regular review of procedures and practices

'Should AI ever be abused by the police, it is essential to have the legal means to punish it exemplarily and to inform the public about it.'
– Slovakia

Enforceability and Redress: *it is possible to compel police to comply with all requests to improve their AI practices*

Participants reflected on Enforceability primarily in the form of organisations that should be responsible for enforcing appropriate AI use, such as an independent oversight body, courts, or judges. Participants were thus mostly concerned with the 'who' instead of the 'what' or 'how' (process, focus, or method).

Respondents also called for a Redress framework. Most frequently, respondents highlighted the need for consequences for any misuse of AI, calling for wrongdoing to be prosecuted and to have legal consequences, both at the organisational level and at the individual level of officers and staff. Calling for additional security measures for AI implementation, participants likewise suggested that the consequences for wrongdoing should be particularly high, including fines, other penalties, and even dismissal. At the same time, some answers highlighted the difficulty for citizens to access the justice system in cases of wrongdoing. This expresses a clear desire for redress procedures that citizens can easily access.

Selected answers referencing Enforceability and Redress	
Enforceability and Redress procedures must ensure	
• The creation of an agency that handles complaints against police use of AI • Compensation procedures for wrongdoing	• Easy and free access to courts and justice • Prosecution of AI misuse with tough consequences: fines, immediate dismissal

Compellability: *it is possible to compel the police to comply with all requests to improve their AI practices*

The Principle of Compellability sits within a higher-level understanding of the Accountability of AI, by which citizens and independent bodies must be able to access information about the implementation and operation of AI by police forces. Therefore, it is not surprising that one of participants' biggest concerns was appropriate independent monitoring of AI use. Furthermore, participants suggested that access to data held about them must be easily accessible, or organisations deploying AI must provide clear information about its use. Open answers thus signal a significant desire for Compellability, although citizen responses indicate that they focus on Compellability primarily in the sense of information provision and thus as an important mechanism to achieve Transparency.

'The use of AI by the police should be auditable by randomly selected groups of people and they should be given all the tools and support to be able to evaluate that use.'
– Republic of Cyprus

Selected answers referencing Compellability	
Compellability must be guaranteed by ensuring that	
• Data held about individuals is provided when requested • Citizens and lawyers have access to data held about individuals	• Mandatory inspections are carried out after the use of AI tools • Randomly selected people audit AI use

Explainability: *all AI practices, systems, and decisions can be fully explained to the public and oversight bodies*

The significance of Explainability emerged in respondents' emphasis on the need for citizens to be informed about police forces' AI use. Open answers specifically called for thorough information about police forces' AI use in such a way that this technology is understandable to lay citizens. This ensures that any person potentially affected by AI, regardless of their level of expertise, is fully aware of its implications and of how and why it is used.

Respondents repeatedly highlighted that explanations of AI use must be simple, '*clear*' and '*complete*'. As one respondent formulated, it is key to '*tell the public exactly what the police can do*'. Yet, responses also show some concerns about potential manipulation in the process of providing information to citizens, suggesting that Explainability must be independently verified to avoid issues such as partial or misleading representation of AI deployment.

'A brochure should be sent out to all homes explaining how and AI is used by the police because I think it will help people to feel safer in the community'
– Ireland

Selected answers referencing Explainability	
Steps to guarantee Explainability	
• Explain AI use in simple terms, including how and why it is being deployed • Explain how AI works and what information it gathers • Explain the benefits of AI	• Make police Commissioner responsible for explaining AI use • Any actions taken need to be both visible and clearly explained • Official letter to be sent out to citizens with detailed explanations

Constructiveness: *police and groups that assess police use of AI always have a constructive attitude in their negotiations with each other*

'Putting oneself in the shoes of the normal population.'
– Germany

Constructiveness found expression in respondents' suggestions to establish a trusting relationship between citizens and police forces, including the opportunity for citizens to be heard in cases of wrongdoing. While a significant number of comments referred to the need for one-directional information to citizens (i.e., Transparency), others imagined this as a more dynamic relationship in which an open dialogue is established between stakeholders, e.g., '*appropriate engagement with various citizen assemblies throughout the formation of policy and implementation*'.

Selected answers referencing Constructiveness	
Constructiveness should be ensured by	
• Officers do not abuse their position of authority and treat citizens fairly • Constructive oversight	• Ongoing dialogue with citizens • Showing empathy towards the positions of citizens

Conduct: *all police uses of AI strictly follow professional standards*

'Be faithful to their service and not abuse their profession.'
– Portugal

Participants had clear proposals for how police should conduct themselves, not only targeting police forces as organisations but also individual officers and staff responsible for AI operations. A core focus was the need to investigate and prosecute incidents of AI misuse while addressing corruption and abuse of power.

One participant summarised this as police should not '*abuse their power and authority in using AI*' while another suggested that individuals responsible for AI use should be regularly replaced. Furthermore, respondents called for additional security protocols that could help reduce potential misconduct, including specialised training and particularly severe legal consequences in cases of AI misuse.

Selected answers referencing Conduct	
Adequate Conduct must be guaranteed by ensuring that	
• Officers abide by the law and serve to protect the public • Employing people of good behaviour and attitude • Officers follow AI guidelines, professional protocols, and the law • Officers focus on prevention rather than punishment • Ethical codes are implemented	• Officers are assessed and adequately trained for AI use • Officers undergo psychological tests and screening prior to AI use • Only experienced officers use AI • Corruption among officers is tackled • Officers operating AI tools are regularly rotated or replaced • AI tools are not used by single officers

Learning Organisation: *Police are continually willing to change their current AI practices based on new knowledge and insights*

The need for continuous learning has been addressed by respondents through calls to perform thorough testing and research about AI, which in turn will help develop AI capabilities that minimise errors before they are deployed. As one respondent noted, it is essential that police '*continue to learn new things and develop*'. Participants have also highlighted the need for continuous training of staff operating AI tools: *'maximum level of education and constant upgrading of knowledge and experience'*.

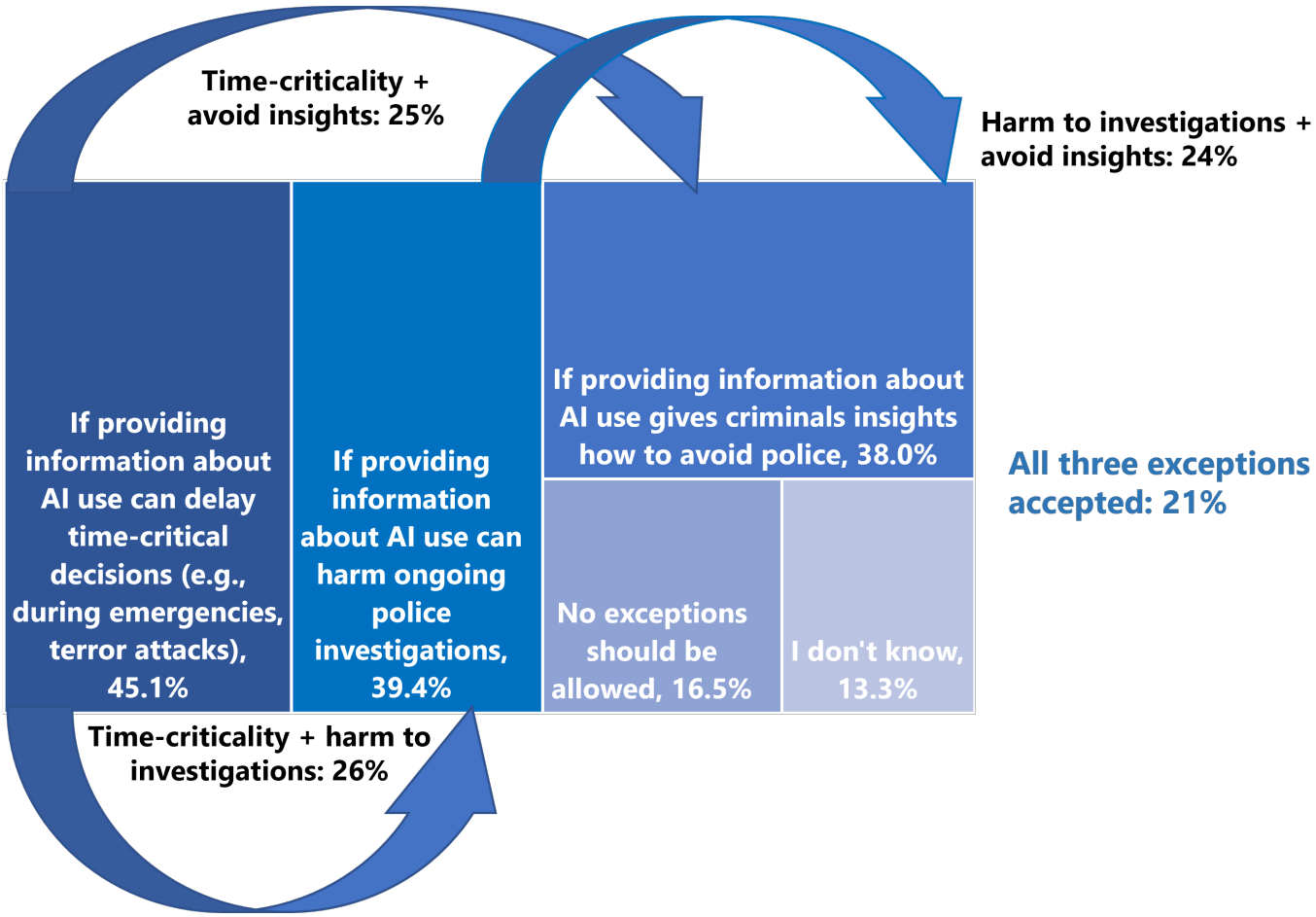
'The use must be constantly evolving and improved by experts.'
– Italy

Selected answers referencing a Learning Organisation	
A Learning Organisation means	
• Regular training for officers operating AI tools • Continuous research to refine procedures • Thorough testing to avoid errors	• Lesson learning to translate into positive change in AI use • Waiting for AI implementation until tested in third countries

Exceptions for the application of AI Accountability

Discussions in the policing and justice domains often reference the need to allow exceptions for police to provide information about their AI capabilities and use. The citizen consultation asked about possible reasons for exceptions to understand whether the public accepts such exceptions to full accountability and, if yes, in which cases.

Only a small percentage of participants (16.5%) refused to permit any exceptions, which suggests that citizens are generally sensitive to the complexity of AI Accountability in policing. The most accepted exceptions were those **due to time-criticality**, i.e., if the request to provide information would delay time-sensitive security decisions. This was followed by the need to **prevent potential harm to ongoing police investigations** and the need to **prevent providing criminals with insights** that could help them avoid the police. About a quarter of participants were willing to grant more than one exception; 20.5% of participants allowed even all three situations as exceptions.

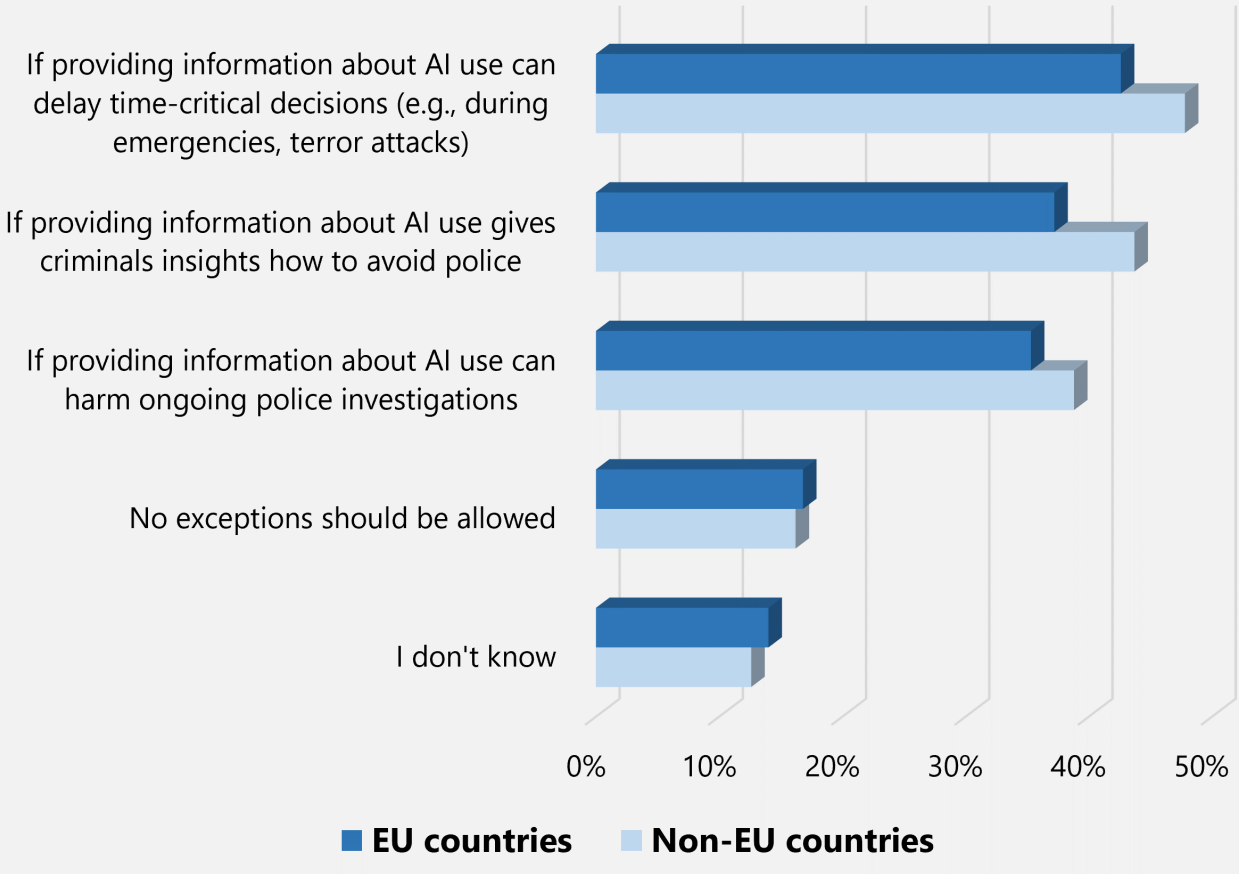


Comparing EU and non-EU countries, the picture is similar for both groups: 16.8% of participants in EU countries and 16.2% in non-EU countries indicated that no exception should be allowed. **Ethnic majority and minority members** did not differ in their overall refusal to grant exceptions (16.5% vs. 17.5%). However, self-described members of an ethnic majority were considerably more willing to grant exceptions due to time-criticality and to prevent criminals from obtaining insights into police operations.

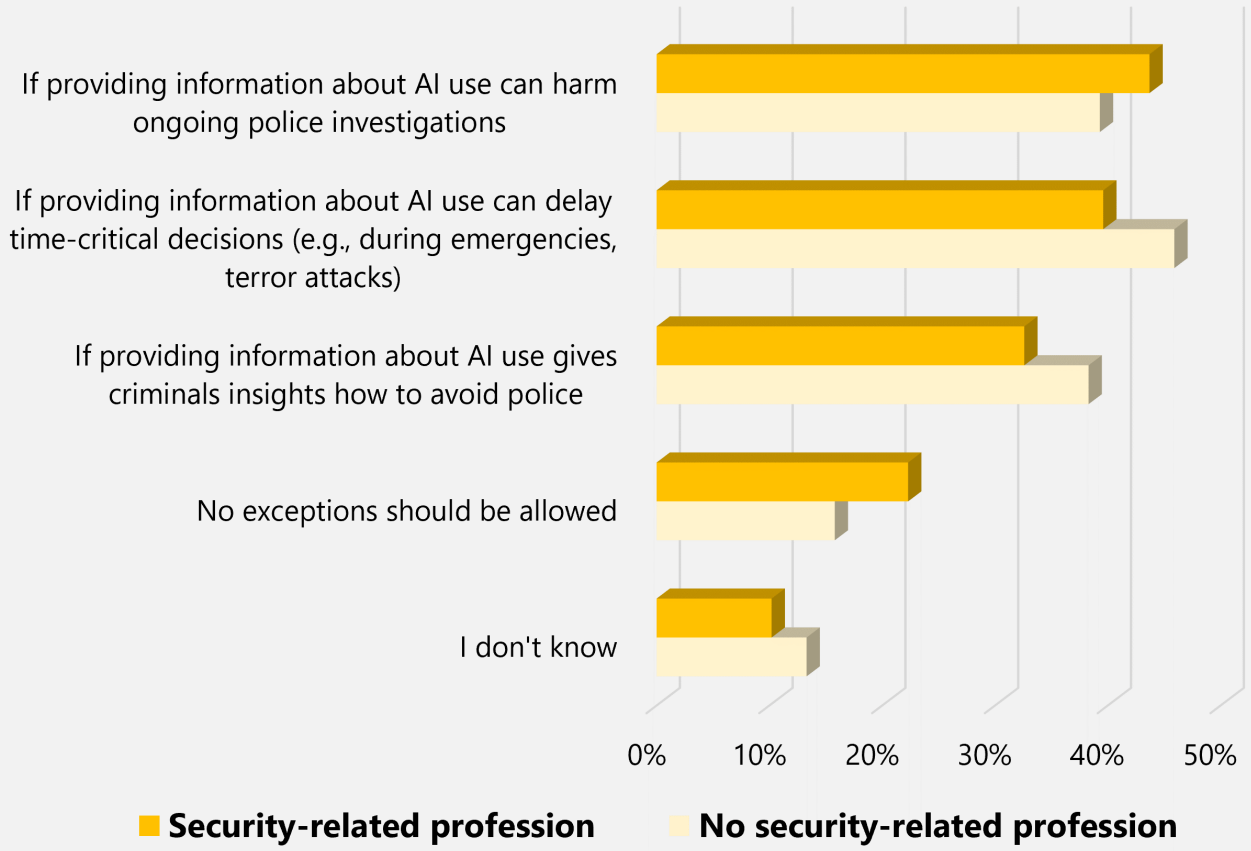
A slightly stronger difference can be observed for **participants with or without a security-related profession**. People with a security-related profession were the least willing to grant exceptions (22.3% compared to 15.8%). They further considered the harm to ongoing police investigations as the most acceptable reason to grant exceptions. In contrast, people working in a non-security-related profession considered time-critical decisions as most relevant.

Additional exceptions mentioned by participants	
• When there is danger to human life or health • For specific cases only, e.g., child abuse • Cases involving private life • Use only if common sense is applied by the police	• Do not use them if they harm citizens • Do not use if rights are violated • Don't use AI at all
Please note: Only 26 participants (0.4% of the sample) chose to add comments, meaning that the information above should not be read as representative of the overall sample.	

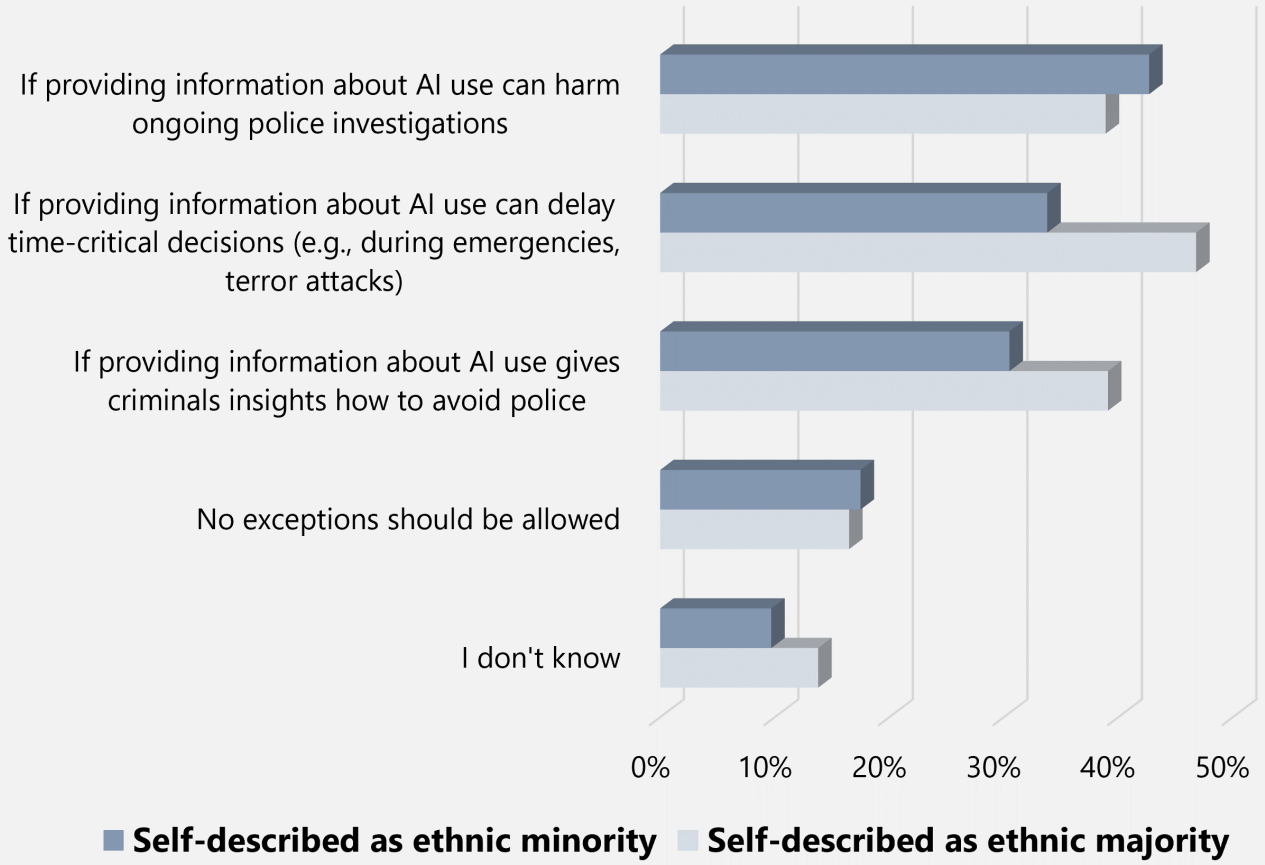
EXCEPTIONS FROM SCRUTINY



EXCEPTIONS FROM SCRUTINY



EXCEPTIONS FROM SCRUTINY



PARTIES RESPONSIBLE FOR AI ACCOUNTABILITY

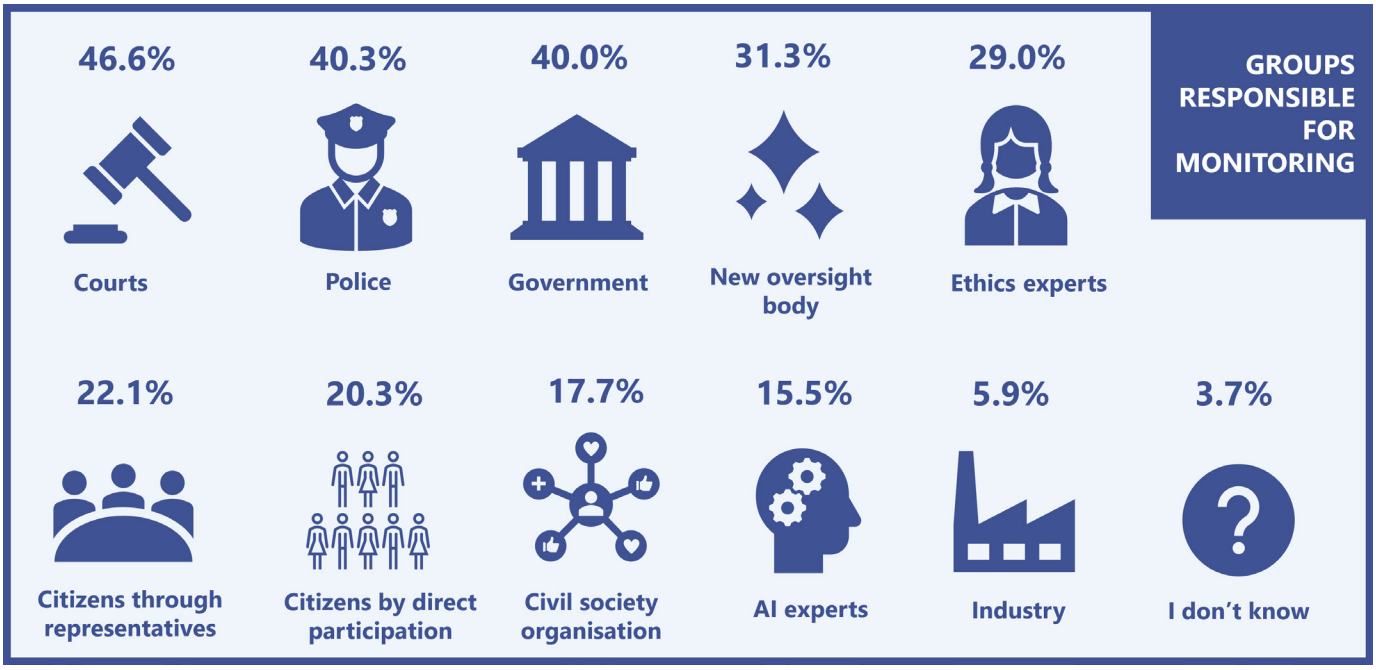
Participants showed clear preferences for the groups and organisations that should be responsible for the monitoring and enforcement of proper AI use, as well as any penalties or redress as part of the AI Accountability process. Answers to this question have direct implications for AI Accountability in that they provide concrete pointers on how to implement and contextualise stakeholder involvement (for instance, who should be involved in AI Accountability conversations: Pluralism principle; who should receive information: Transparency principle; who should be able to compel changes: Compellability principle, etc).

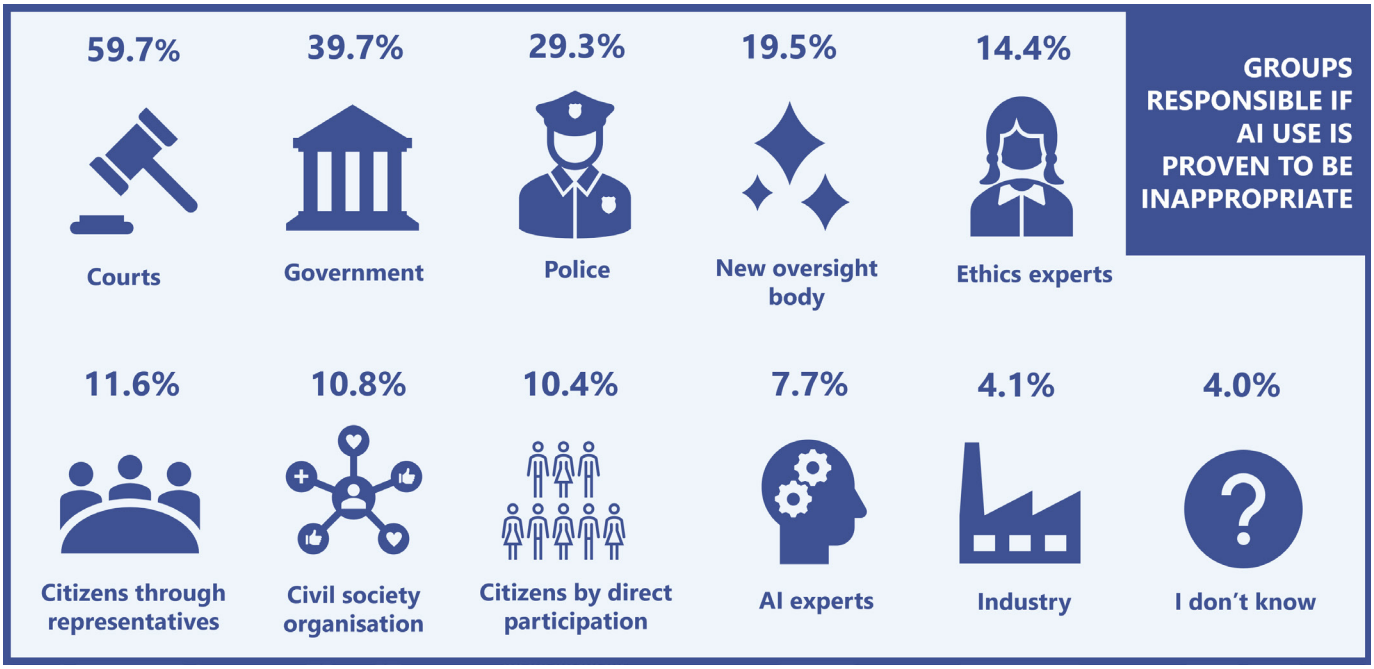
Courts emerged as the preferred bodies for both areas, followed by the police themselves and government ministries. That **police** emerged as an important responsible party is an interesting observation, as it means that a large number of citizens expect police to play an active part in the AI Accountability process; although more for monitoring than for the enforcement of corrections and penalties. A third of participants preferred to establish a **new oversight body**.

Interestingly, only a relatively small proportion of participants called on **citizens** to be part of the accountability process, either through direct participation or through representation. Especially for enforcement, i.e., in the case that AI use is proven to be inappropriate, citizens were only considered by approx. 11% of participants. **Industry** participation was seen as the least important.

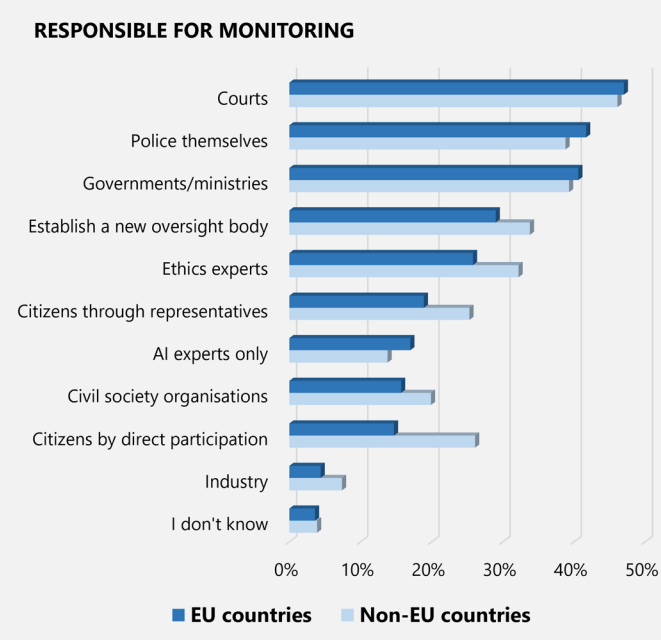
Participants were asked to provide additional groups they considered relevant. 115 participants took up this offer to add additional groups or actors. Amongst them were proposals for the inclusion of journalists, randomly chosen citizens, third-country experts, or even the option for AI to self-monitor, providing pointers for possible expansion of the accountability landscape.

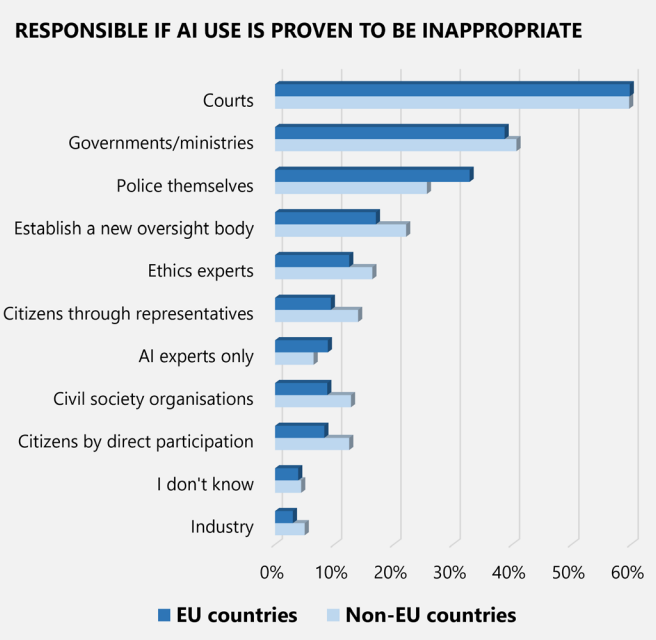
Selected groups mentioned by participants	
• Independent AI experts • Independent oversight body • NGOs and human rights organisations • Body of randomly chosen citizens • Taxpayers • Government bodies • Ethics experts	• Third-country experts • Journalists • Members of the legal system • Police-related: e.g., internal affairs, crime commissioners • Psychologists and psychiatrists • Religious organisations • Use AI to self-monitor AI
Please note: Only 115 participants (1.7%) chose to add additional groups for inclusion, meaning that the comments above should not be read as representative of the overall sample.	

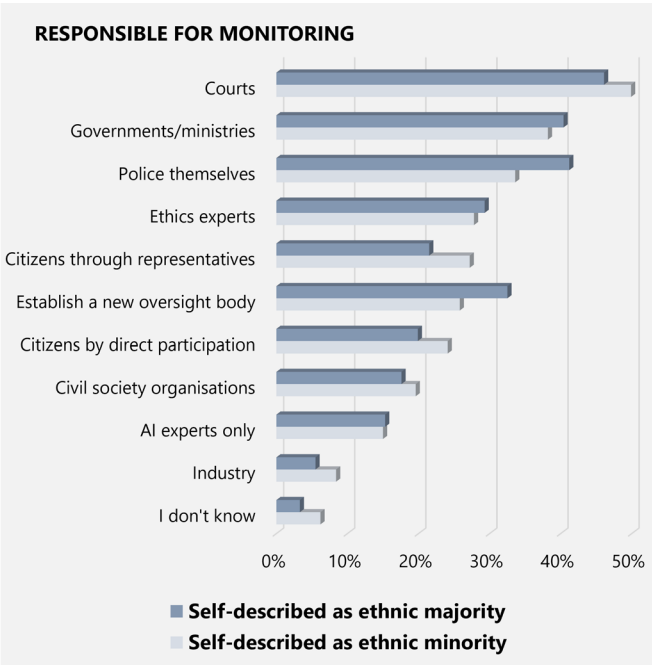


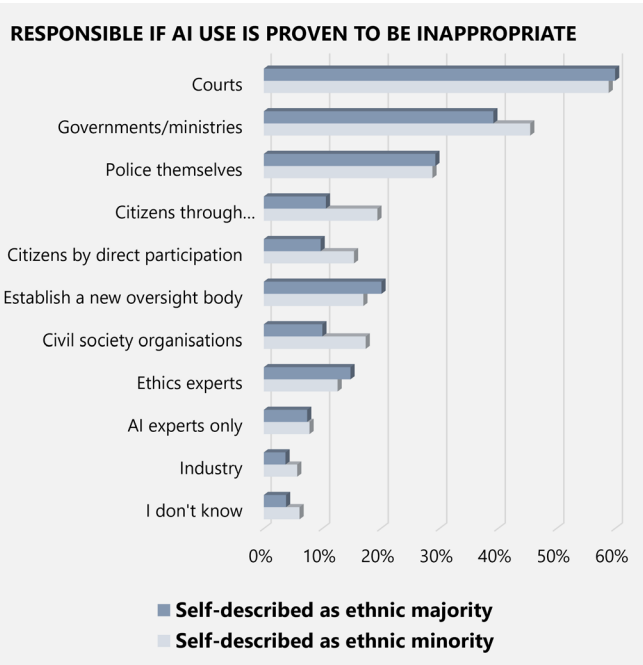


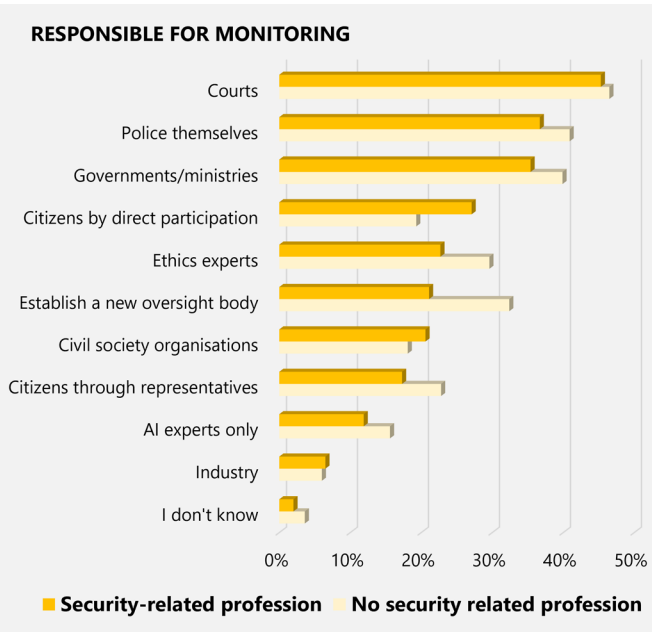
Participants from **EU and non-EU countries** were mostly in agreement about the types of groups and organisations that should be responsible for monitoring and enforcement. The main difference can be seen in a somewhat stronger appreciation of citizen participation in non-EU countries. The same is true for people in **security-related professions** versus non-security-related professions; the main difference here was the greater role of courts in cases where AI is proven to be used inappropriately.

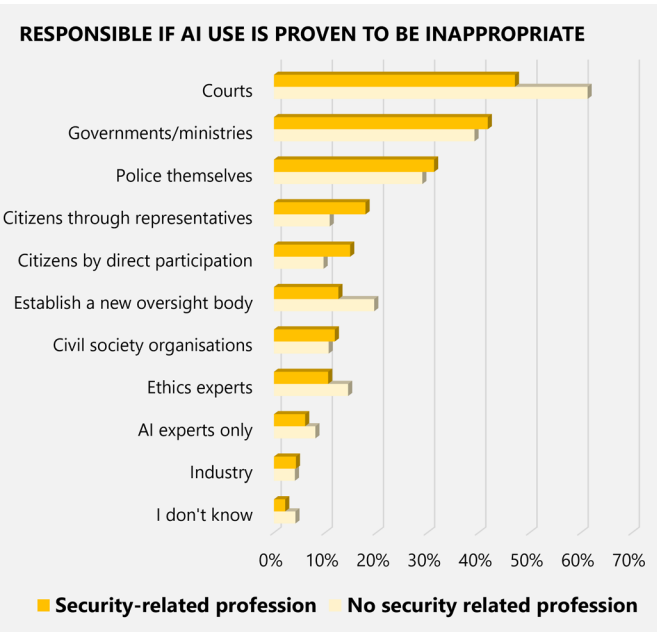

RESPONSIBLE FOR MONITORING
Courts
Police themselves
Governments/ministries
Establish a new oversight body
Ethics experts
Citizens through representatives
AI experts only
Civil society organisations
Citizens by direct participation
Industry
I don't know
0%
10%
20%
30%
40%
50%
EU countries
Non-EU countries


RESPONSIBLE IF AI USE IS PROVEN TO BE INAPPROPRIATE
Courts
Governments/ministries
Police themselves
Establish a new oversight body
Ethics experts
Citizens through representatives
AI experts only
Civil society organisations
Citizens by direct participation
I don't know
Industry
0%
10%
20%
30%
40%
50%
60%
EU countries
Non-EU countries

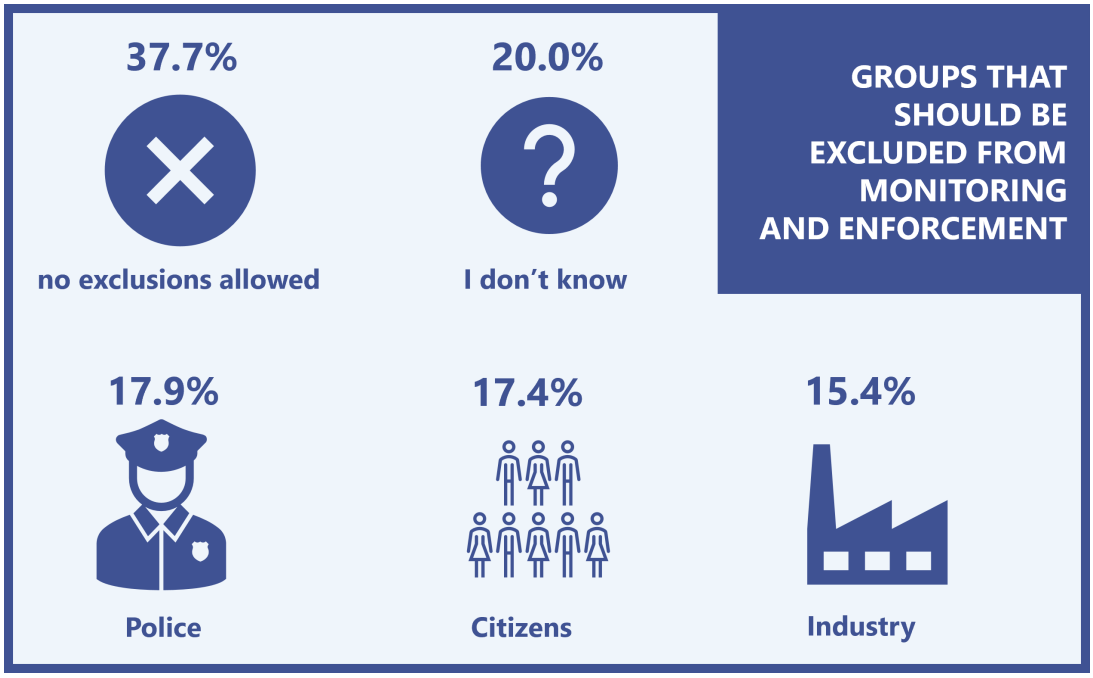

RESPONSIBLE FOR MONITORING
Courts
Governments/ministries
Police themselves
Ethics experts
Citizens through representatives
Establish a new oversight body
Citizens by direct participation
Civil society organisations
AI experts only
Industry
I don't know
0%
10%
20%
30%
40%
50%
Self-described as ethnic majority
Self-described as ethnic minority


RESPONSIBLE IF AI USE IS PROVEN TO BE INAPPROPRIATE
Courts
Governments/ministries
Police themselves
Citizens through...
Citizens by direct participation
Establish a new oversight body
Civil society organisations
Ethics experts
AI experts only
Industry
I don't know
0%
10%
20%
30%
40%
50%
60%
Self-described as ethnic majority
Self-described as ethnic minority


RESPONSIBLE FOR MONITORING
Courts
Police themselves
Governments/ministries
Citizens by direct participation
Ethics experts
Establish a new oversight body
Civil society organisations
Citizens through representatives
AI experts only
Industry
I don't know
0%
10%
20%
30%
40%
50%
Security-related profession
No security related profession


RESPONSIBLE IF AI USE IS PROVEN TO BE INAPPROPRIATE
Courts
Governments/ministries
Police themselves
Citizens through representatives
Citizens by direct participation
Establish a new oversight body
Civil society organisations
Ethics experts
AI experts only
Industry
I don't know
0%
10%
20%
30%
40%
50%
60%
70%
Security-related profession
No security related profession

When asked for the **explicit exclusion of specific groups**, nearly 40% of participants indicated that no exclusions should be allowed. Only 18% of participants preferred to exclude police from monitoring and enforcing their appropriate AI use, which is at a similar level to citizens and industry. Additional groups were identified in the open answer option; key amongst them is the exclusion of '*governments*', '*politicians*', and '*criminals*'.



Selected groups for exclusion mentioned by participants	
• Government and politics-related, including councils, government representatives, ministers, political parties, politicians • Children • The general population • Private industry • Criminals	• LEAs, secret services and the army • Anyone not vetted • Groups that advocate for special interests, such as societal, political or religious groups • Unrelated third parties • If exclusions are granted, these should be formally approved
Please note: Only 72 participants (1.1%) chose to add additional groups for exclusion, meaning that the comments above should not be read as representative of the overall sample.	

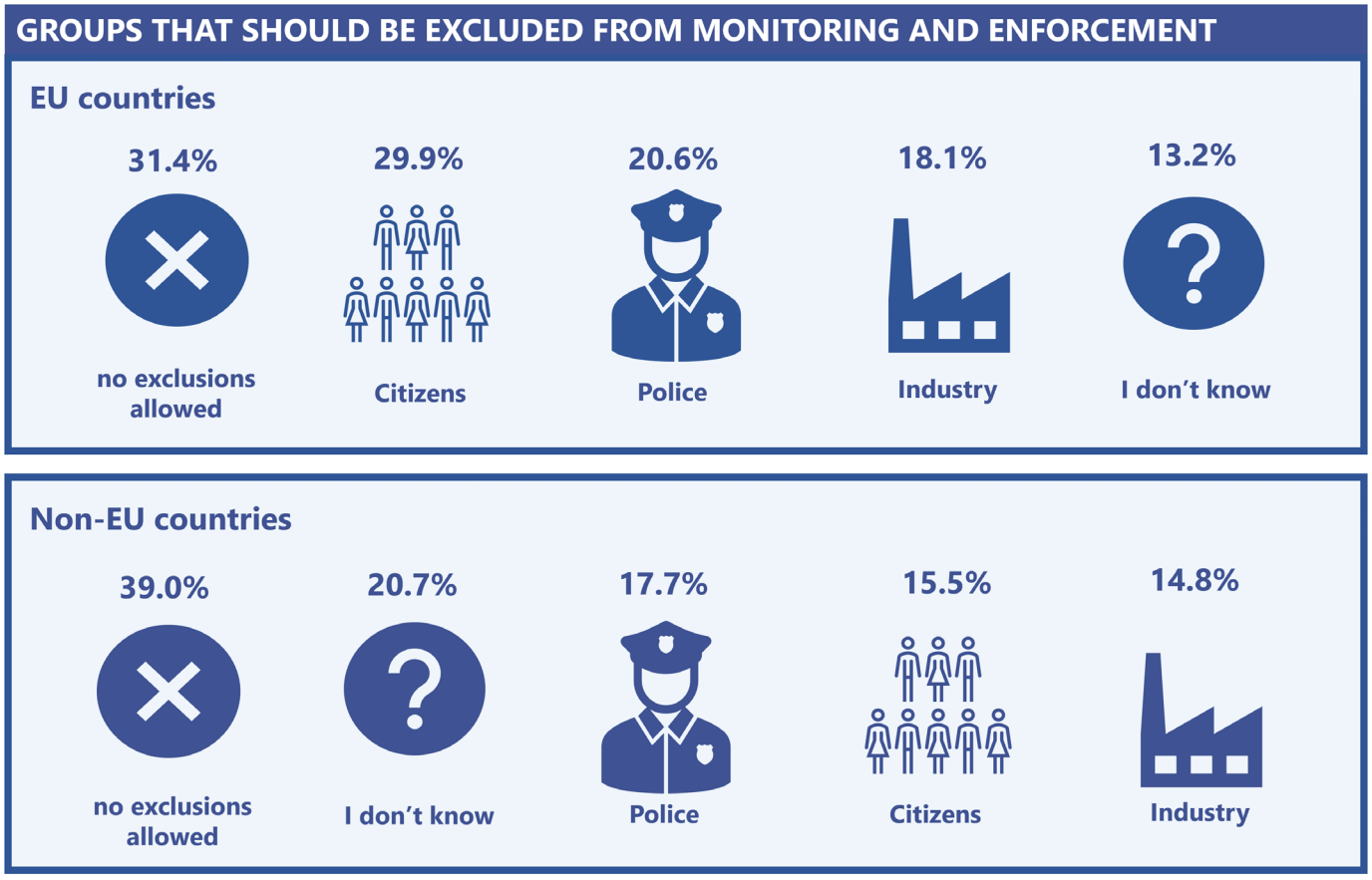
Some of the additional groups proposed in the open answers warrant further consideration. It may be problematic, for instance, to exclude children by default, as they may be equally affected by AI as adults. The same could be said about most other groups listed by participants.

The individual answers illustrate that citizen perspectives are likely to depend on context and personal experience and that communities may differ in which groups they would like to see included or excluded. Such differences and sensitivities will be important to take into account when engaging with citizens and diverse communities in the AI Accountability process.

What the list of excluded groups also demonstrates is that the public may hold certain preconceived expectations about who is seen as 'qualified' or 'acceptable' to decide about AI use by the police. Such expectations may affect how communities perceive efforts by police or other stakeholders in the internal security domain to engage with multiple stakeholders (e.g., in aiming to fulfil the Universality principle). Multi-stakeholder engagement may thus cause tensions, which will have to be expected, planned for, and managed sensitively to ensure desirable outcomes.

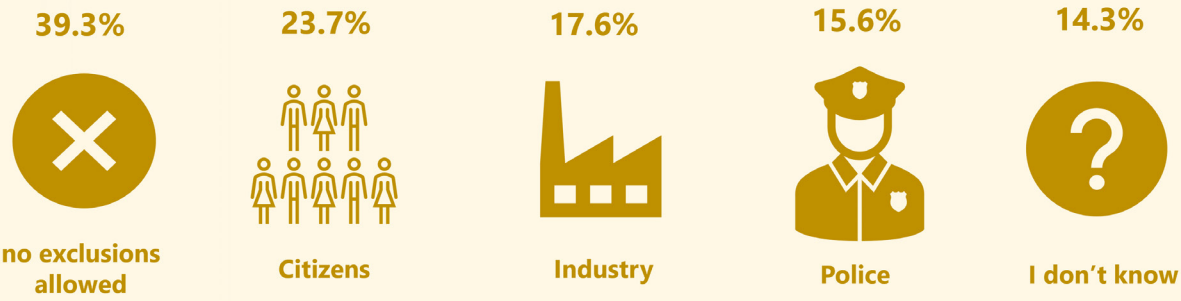
Group differences

Comparing subgroups in **EU and non-EU countries** yielded very similar results, although we found a slightly higher prevalence in non-EU countries to exclude police. Participants with a **security-related profession** tended to be more negative about including citizens compared to people not working in security-related professions. Importantly, of all groups, **people identifying as ethnic minorities** were the most likely to prefer the exclusion of citizens and police from monitoring AI Accountability and redress actions.

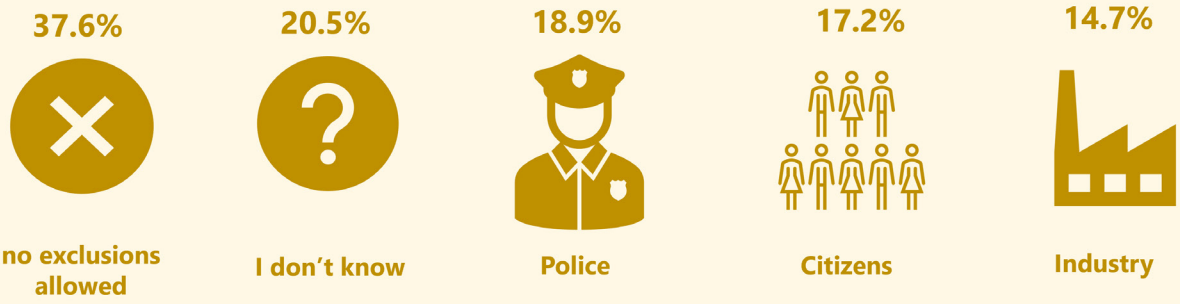


GROUPS THAT SHOULD BE EXCLUDED FROM MONITORING AND ENFORCEMENT

Security-related profession



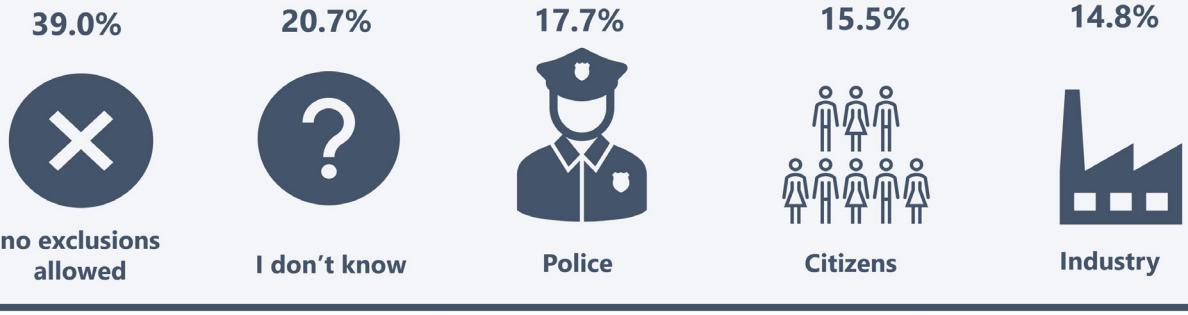
No security-related profession



GROUPS THAT SHOULD BE EXCLUDED FROM MONITORING AND ENFORCEMENT

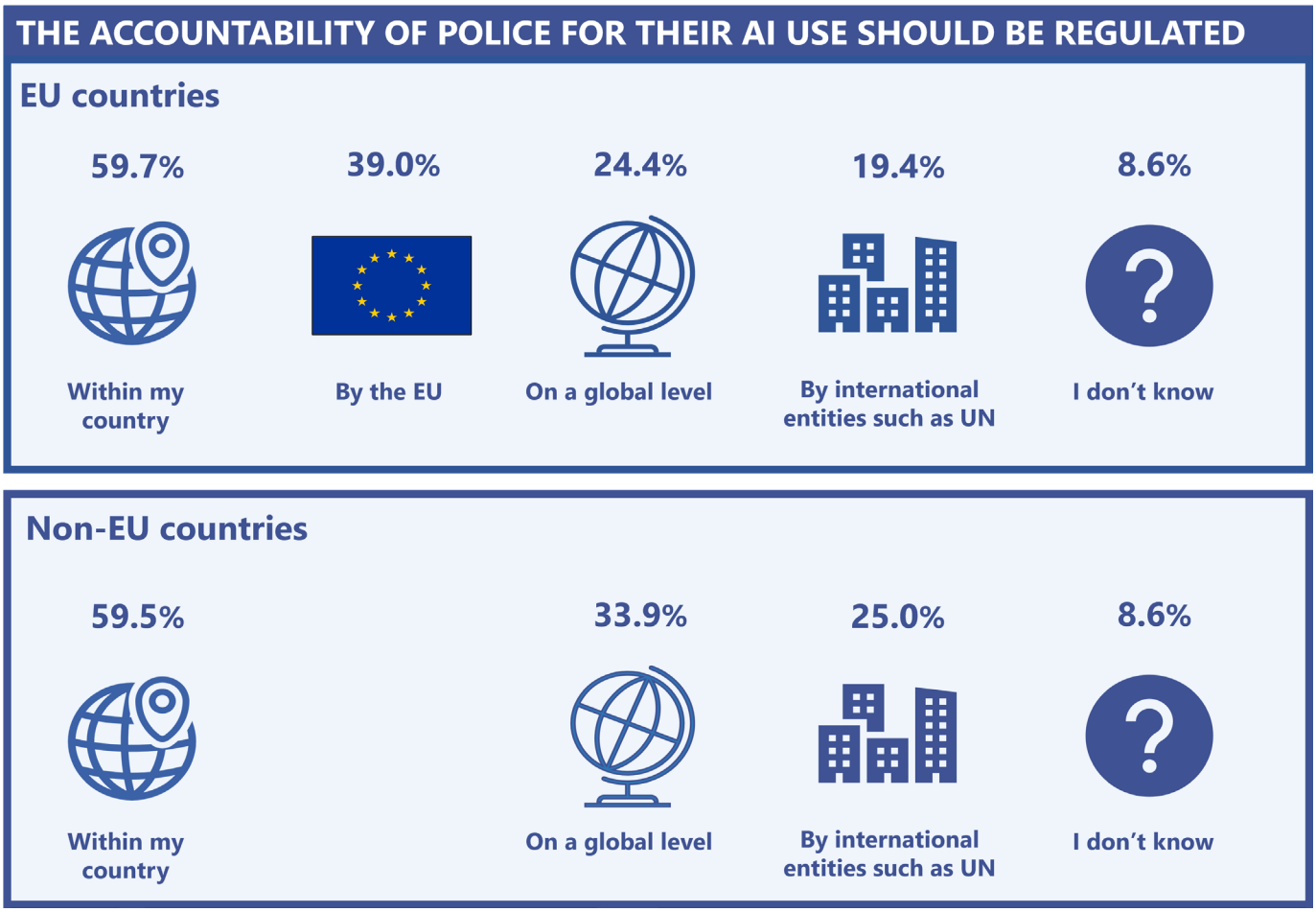
Self-described as ethnic minority

Self-described as ethnic majority



PREFERRED REGULATION LEVEL

Citizens had a clear preference for regulation at the national level. This was followed by regulation at the global level and, to a lesser extent, by international organisations. Participants in EU and non-EU countries did not show marked differences in these preferences. Participants in EU countries were also asked specifically for regulation by the EU, to which 39% agreed.



SUMMARY

The AP4AI citizen consultation offers unique insights not only into the generic questions of concerns or benefits but also into very concrete expectations the public holds about AI use by the police—and specifically how AI use should be regulated and in which way police should be held accountable.

Below, we summarise the main findings:

- **Areas with support for AI use:** 67% of participants agreed/strongly agreed that AI can greatly profit society (8% disagreed/strongly disagreed); 87.6% agreed/strongly agreed that AI should be used for the protection of children and vulnerable groups; 83.0% agreed/strongly agreed that AI should be used to detect criminals and criminal organisations; 73.4% agreed/strongly agreed that AI should be used to predict crimes before they happen.

- **Concerns about AI use:** 51% of participants agree that police in their country make sufficient efforts to avoid the negative consequences of AI; 60% agree that police respect citizens' rights in their activities; 37.2% had concerns about potential negative effects due to police decisions based on AI, while 48.7% noted concerns about police using AI to monitor their online information and 46.7% had concerns about police using AI to monitor their offline activities.

- **Need for AI Accountability:** 90% of participants agreed that police should be held to account for the way they use AI, 88.3% for the consequences of their AI use; 81.8% rated the existence of an overarching AI Accountability Framework as important or very important; all 12 AP4AI Principles emerged as important or very important.

- **Acceptance of current AI Accountability mechanisms:** 31.9% of participants considered existing accountability mechanisms adequate, 25.8% perceived them as too weak, 8.1% rated them as too restrictive, and 34.2% were unable to judge ('I don't know'); perceptions of current AI Accountability mechanisms were affected by level of AI expertise

- **Groups expected to oversee AI use by the police:** main groups considered responsible for monitoring AI use by police: 46.6% courts, 40.3% police, 40.0% governments; citizen involvement: 21.1% wished citizens to participate through representatives, 20.3% by direct participation; main groups considered responsible if AI use is proven to be inappropriate: 59.7% courts, 39.7% governments, and 29.3% police; citizen involvement: 11.6% wished citizens to participate through representatives, 10.6% by direct participation.

- **Level at which AI Accountability should be regulated:** for EU citizens: 59.7% of EU participants preferred regulation within their own country, 39.0% by the EU, 24.4% on a global level, 19.4% by international organisations such as the UN; for non-EU citizens: 59.5% within their own country, 33.9% on a global level, 25.0% by international entities such as the UN.

CITIZEN-BASED RECOMMENDATIONS TO IMPLEMENT AI ACCOUNTABILITY

Citizens across the 30 countries expressed a broad range of expectations and suggestions on **how to implement AI Accountability**, as well as **how to create and maintain their trust** in AI use by police forces. These expectations and suggestions address disparate aspects of the AI landscape, ranging from AI systems, tools, and data to expectations about laws and regulations and ways to ensure awareness and continuous learning. Below, we summarised these entries into concrete recommendations for the implementation of AI Accountability.

AI SYSTEMS AND TOOLS

Regularly **Assessing and Updating** AI systems:

- Regularly **audit AI systems** to identify and rectify any biases, errors, or potential risks. This will help maintain the integrity of the tools and ensure their use remains responsible and beneficial to society.
- Prioritise **continuous improvement by** investing in research and actively incorporating best practices into existing AI systems. This ensures that the tools remain efficient, secure, and beneficial to society.

Promoting **Accountable and Transparent Use of AI**:

- Establish **clear guidelines for public disclosure** that require the provision of transparent information to the public regarding the use of AI. This includes disclosing details about the data sources and algorithms/models utilised, as well as the decision-making processes involved. By setting clear expectations, stakeholders can have access to essential information necessary for understanding the implications of AI systems and tools in policing.
- Publish regular **transparency reports** on system performance, biases, and actions taken to address issues, while proactively disclosing updates or modifications.

DATA

Protecting **Privacy and Data Security**:

- Implement **strong encryption, access controls, and secure data storage practices** to prevent unauthorised access, breaches, or misuse of personal information.
- **Properly anonymise data and limit its storage duration** to further safeguard citizen privacy rights, thereby maintaining public trust in the responsible handling of personal information.
- Ensure **compliance with applicable data protection laws and regulations** throughout the lifecycle of AI tools.

LAWS AND REGULATIONS

Establishing a **Robust Legal Framework**:

- Develop **comprehensive legislation** specifically addressing AI use by the police. This should define clear guidelines and requirements for data collection, retention, and access.
- Regularly **review and update** the legal framework to keep pace with technological advancements.
- Implement **legal safeguards to protect citizens' privacy and prevent misuse** of AI technologies.
- **Establish how citizens will have access to information about AI use and data** collected about individuals.
- Ensure the framework encompasses **accountability measures**, including the identification of bodies or organisations responsible for independent oversight.
- **Involve legal and AI experts**, policymakers, citizens, and other stakeholders in the development of the framework.

RISK ASSESSMENT AND MANAGEMENT

Mitigating **Discrimination Risks:**

- Avoid unfair treatment of citizens in the use of AI. This requires **thorough testing, evaluation, and ongoing monitoring of AI systems** to detect and address any biases or disproportionate impacts that may arise.
- By **prioritising fairness and equity**, the potential for discriminatory outcomes can be minimised, promoting trust and confidence in AI use.
- **Recognise and mitigate systemic biases and historical discrimination** faced by certain populations, including women, migrants, disabled people, and ethnic, religious, LGBTQI+, and indigenous communities. By prioritising the needs and rights of marginalised and vulnerable communities, the potential for AI to exacerbate discrimination can be mitigated, fostering equal treatment and protection for all citizens.
- Allow the **operation of AI systems to fully trained staff only**. This reduces the risk of misuse and misinterpretation and ensures AI tools are used to their full potential while respecting legal and ethical standard.

- **Promote diversity to ensure that the staff operating AI tools** reflects the heterogeneity among citizens potentially affected by AI deployment. A diverse team brings a broad range of perspectives and experiences, contributing to fairness, inclusivity, and representation.

OVERSIGHT AND REDRESS PROCESS

Establishing **Independent Oversight**:

- Establish an **independent body to oversee and regulate the use of AI in policing** and ensure compliance with ethical and legal standards. This body should be responsible for setting ethical guidelines, monitoring AI systems' usage, and ensuring adherence to legal standards. By providing independent oversight, public trust can be enhanced, and potential risks or abuses associated with AI deployment can be effectively addressed.
- **Regularly audit and evaluate AI systems** to assess their effectiveness, fairness, and adherence to guidelines. Regular audits can identify any biases, errors, or shortcomings in the systems and allow for necessary improvements.
- **Allow oversight of civil society organisations** that work to protect citizens' rights. By involving them in the oversight process, diverse perspectives can be incorporated, offering valuable insights, expertise, and independent assessments.

Creating a **Redress System**:

- Develop a user-friendly online **platform for reporting incidents of AI misuse and obtaining support**. The reporting system should allow citizens to submit their concerns regarding potential violations and enable redress procedures. A dedicated support team should provide citizens with free, timely assistance, both online and offline.
- **Collaborate with relevant organisations for investigations**, including legal and advocacy groups that can offer additional support to affected citizens. These collaborations can provide expertise, independent assessments and guidance throughout the investigative process, ensuring that citizen' rights are protected and their concerns are addressed appropriately.
- It is essential that the **privacy and confidentiality of individuals who report incidents is protected** to encourage their participation in the process.
- Conduct **public awareness campaigns** on reporting options and citizen rights. These campaigns play a vital role in empowering individuals to understand their rights, reporting options, and the available mechanisms for addressing incidents of AI misuse. These campaigns should aim to educate citizens about the importance of reporting, the potential impacts of AI misuse and the procedures for filing complaints.

ROBUST ACCOUNTABILITY EVIDENCE

Establishing a Framework for **Individual Conduct Accountability**:

- Define **clear protocols for AI tool usage and monitor staff performance** to maintain adherence to these standards. Perform regular audits of staff compliance to identify potential issues and validate the effectiveness of protocols.
- Enforce **disciplinary measures** to address any misconduct.
- Build **robust mechanisms for reporting concerns** or violations to encourage transparency and trust within the organisation.
- Offer **confidential support channels for staff navigating ethical dilemmas**, promoting a safe space for dialogue and ethical decision-making.

STAKEHOLDERS

Ensuring **Multi-Stakeholder Collaboration**:

- Establish mechanisms for **involving diverse stakeholders**, including civil society organisations, academic experts from both technical and social sciences, legal professionals, and affected communities, in the development, implementation, and evaluation of AI systems.
- Promote the **inclusion of impacted communities and marginalised groups** in decision-making processes to ensure that their perspectives and needs are taken into account when designing and implementing AI systems, promoting greater trust and legitimacy in police practices.

Fostering **Community Engagement** and Consultation:

- **Involve citizens and community representatives** in decision-making processes regarding AI implementation by facilitating platforms for dialogue and consultation. This can be achieved by establishing mechanisms such as citizen advisory boards, consultations, or public forums where individuals can participate, provide input, and contribute their perspectives. By involving the community in these processes, their concerns and needs can be better understood and included, leading to more inclusive and accountable AI deployment.
- Conduct **public awareness campaigns** and educational initiatives to promote understanding of AI technologies. These initiatives should aim to explain the capabilities, limitations, and potential impacts of AI in a clear and accessible manner.

AWARENESS AND LEARNING

Provide ongoing **Training and Education**:

- Conduct **repeated public awareness campaigns** about the capabilities, limitations, and potential risks of AI systems in policing, promoting transparency and accountability.
- Facilitate **ongoing professional development** programmes focused on AI best practices to ensure staff acquire up-to-date knowledge and skills, enhancing their capacity to use AI tools responsibly and effectively.
- **Train police personnel** on the ethical use of AI, including awareness of potential biases and the appropriate interpretation of AI-generated outputs.
- **Regularly update training content** to reflect advances in AI technology, preparing staff to adapt to the evolving technological landscape.
- Promote a culture of accountable AI use through **continuous learning and development** programs. These should not be a one-time initiative but an ongoing effort to keep pace with the rapidly evolving AI landscape, ensuring accountable use of AI.

ENDNOTES AND REFERENCES

1 E.g., Pew Research Center. (March 2022). AI and Human Enhancement: Americans' Openness Is Tempered by a Range of Concerns. https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2022/03/PS_2022.03.17_AI-HE_REPORT.pdf

2 Carrasco, M., Mills, S., Whybrew, A., & Jura, A. (2019). The Citizen's Perspective on the Use of AI in Government. BCG Digital Government Benchmarking. https://www.bcg.com/publications/2019/citizen-perspective-use-artificial-intelligence-government-digital-benchmarking

3 Akhgar, B., Bayerl, P.S., Bailey, K., Dennis, R., Heyes, S., Lyle, A., Raven, A., Sampson, F., & Gercke, M. (2022). AP4AI Report on Expert Consultations. AP4AI Report 1. https://www.ap4ai.eu

4 Haataja, M., van de Fliert, L., & Rautio, P. (2020). Public AI Registers: Realising AI Transparency and Civic Participation in Government Use of AI. https://openresearch.amsterdam/nl/page/73074/public-ai-registers

5 Ada Lovelace Institute, AI Now Institute and Open Government Partnership. (2021). Algorithmic Accountability for the Public Sector. https://www.opengovpartnership.org/documents/algorithmic-accountability-public-sector

6 IEEE. (2019). Ethically Aligned Design. https://standards.ieee.org/wp-content/uploads/import/documents/other/ead1e.pdf

7 Alan Turing Institute. (2021). AI Strategy Survey Results. https://www.turing.ac.uk/sites/default/files/2021-09/ai-strategy-survey_results_020921.pdf

8 Committee of Standards in Public Life. (2020). Artificial Intelligence and Public Standards. A Review by the Committee on Standards in Public Life. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/868284/Web_Version_AI_and_Public_Standards.PDF

9 Akhgar, B., Bayerl, P.S., Bailey, K., Dennis, R., Gibson, H., Heyes, S., Lyle, A., Raven, A., & Sampson, F. (2022). AP4AI Framework Blueprint. AP4AI Report 2. Available at: https://www.ap4ai.eu

10 Akhgar, B., Bayerl, P.S., Bailey, K., Dennis, R., Heyes, S., Lyle, A., Raven, A., Sampson, F., & Gercke, M. (2022). AP4AI Report on Expert Consultations. AP4AI Report 1. Available at: https://www.ap4ai.eu

11 https://population.un.org/wpp/

PROJECT COORDINATION

CENTRIC (Centre of Excellence for Terrorism, Resilience, Intelligence and Organised Crime Research): CENTRIC is a multi-disciplinary and end-user focused centre of excellence for end-user driven innovations in the field of security. The global reach of CENTRIC links both academic and professional expertise across a range of disciplines providing unique opportunities to progress ground-breaking research. The mission of CENTRIC is to provide a platform for researchers, practitioners, policy makers and the public to focus on applied security research. CENTRIC is a publicly funded organisation.

Europol Innovation Lab: The Europol Innovation Lab supports the European law enforcement community in the area of innovation. The Lab provides a structure and a set of services to law enforcement agencies to avoid duplication of work, create synergies and pool resources in order to co-create innovative solutions which can boost operational work of investigators and analysts.

CONTACT

Accountability Principles for Artificial Intelligence (AP4AI)

Website: www.ap4ai.eu
LinkedIn: AP4AI
X: @AP4AI_project
Email: CENTRIC@shu.ac.uk; Innovation-Lab@europol.europe.eu