# AI Scrutiny Panel

## MINUTES

| Date | Time | Venue |
|---|---|---|
| 4th November 2026 | 18:30 | MS TEAMS |

| | Item |
|---|---|
| 1. | **Welcome and apologies** |
| | **Attendees:** [redacted] |
| | The Chair welcomed attendees and set out the purpose of the panel. Introductions were given by all in attendance.<br><br>No apologies were received.<br><br>The panel's role is to provide independent scrutiny and constructive challenge regarding AI deployment within the Constabulary. The focus is not on operational or procurement decisions, but on whether safeguards, governance arrangements and public protections are operating effectively in practice.<br><br>Chief Superintendent Sherrie Nash thanked panel members and highlighted that the panel is believed to be among the first external AI scrutiny panels in the UK. The forum is intended to complement internal AI governance and ethics arrangements.<br><br>MP gave a reminder to all present of the need to comply with confidentiality agreements and the expectation to be participating from a private space. |

| | |
|---|---|
| **2.** | **Declarations of interest** *(Standing agenda item)* |
| | None received |
| **3.** | **Approve minutes of previous meetings** *(Standing agenda item)* |
| | N/A – this was the first meeting of the Panel |
| **4.** | **Introduction to Citizens Assist & Example of Operation Use** |
| | Chief Superintendent Sherrie Nash gave an introduction to Sgt Richard Diggle who then presented Citizens Assist to panel members. Members had already received an information pack ahead of the meeting, so the key points that were highlighted at this stage were:<br><br><u>Context</u><br>- Citizen Assist supports the Constabulary's wider digital transformation ambitions.<br>- The tool is intended to improve public access to policing services through an additional reporting channel.<br>- It is not designed to replace existing methods of contact such as: 999, 101, Police station attendance, Existing online forms.<br>- The primary aim is to improve efficiency whilst maintaining public trust and confidence.<br><br><u>Overview of Pilot</u><br>- Approximately one-third of crime demand arrives through online reporting.<br>- Citizen Assist is an AI-enabled conversational reporting tool integrated into Single Online Home.<br>- The system currently covers selected high-volume, lower-risk crime types.<br>- The AI: Collects information, Asks clarifying questions, Directs users to appropriate reporting routes, Does not make investigative, safeguarding or crime-recording decisions<br>- All reports continue to undergo human review.<br><br><u>Performance Data (Pilot)</u><br>Since implementation:<br>- Over 1,000 conversations have taken place.<br>- 487 submissions have been completed. |

| | |
|---|---|
| | - Around 13% of eligible reports are now being submitted through Citizen Assist.<br>- A single technical submission failure occurred and was rectified.<br><br>Most non-completed journeys related to:<br>- Redirection to alternative reporting routes.<br>- Users abandoning reports.<br>- Emergency referrals. |
| **5.** | **Questions from Panel on tech** |
| | **Q: Would members of the public always understand they were interacting with AI and would some users mistakenly believe they were communicating with a human operator?**<br>Response: The tool identifies itself as an AI assistant, users are repeatedly informed they are not speaking to a police employee, existing reporting methods remain available.<br><br>Panel members highlighted the need to consider users with limited digital literacy or understanding of AI concepts. |
| | **Q: Could AI identify safeguarding issues which may not be explicitly disclosed, such as domestic abuse indicators embedded within a wider narrative.**<br>Response: Citizen Assist identifies keywords, relationships and contextual indicators. The system is designed to recognise potential domestic relationships and other vulnerability markers. Human review remains mandatory and provides an additional safeguard.<br><br>The panel observed that AI may potentially provide richer contextual information than traditional forms by encouraging a more conversational narrative. |
| | **Q: Has it been considered that individuals may be more willing to make exaggerated or false allegations when interacting with AI compared with a human?**<br>Response: The Force acknowledged this is difficult to measure currently and indicated this may become an area for future analysis once sufficient data exists.<br><br>Panel members explored whether considerations have been given to:<br>- Hallucinations |

| | |
|---|---|
| | - Prompt manipulation<br>- Attempts to make the AI perform unintended tasks<br>- Cyber security risks<br>- System resilience<br><br>The Force confirmed:<br>- Extensive testing had been completed before launch.<br>- Public users were included in testing.<br>- Hallucinations and failures identified during testing were fed back for correction.<br>- Guardrails are in place to prevent the AI undertaking unrelated tasks.<br>- Ongoing monitoring of performance forms part of the pilot evaluation.<br><br>Members expressed an interest in undertaking controlled testing of the system's resilience and safeguards. |
| | Clarification was sought regarding:<br>- Technical architecture<br>- Hosting arrangements<br>- Governance arrangements<br>- Evaluation methodology<br>- Success criteria<br><br>The Force confirmed:<br>- The product is hosted within police-approved infrastructure.<br>- Information security, DPIA and AI assurance processes have been completed.<br>- National AI governance, CPS processes and Police AI governance boards have reviewed the product.<br>- Evaluation is being undertaken by the Change & Research Team.<br>- Benefits, risks, ethics and operational impacts are being monitored.<br><br>Members stressed the importance of robust data collection to evidence effectiveness and support future decision-making. |

| | |
|---|---|
| | **Q: Could information be lost, altered or insufficiently represented as AI converts conversations into reports?**<br>Response: Full transcripts are retained and human review of transcripts remains mandatory. Earlier testing identified occasions where important detail was not sufficiently captured and this informed subsequent improvements. Ongoing review continues during the pilot.<br><br>Retention of full transcripts was viewed as a critical safeguard and means of maintaining evidential integrity. |
| | Members explored:<br>- Public perception of AI<br>- Potential bias<br>- Disproportionality<br>- Fairness in outcomes<br><br>The Force confirmed:<br>- Bias testing had been completed using diverse participants.<br>- No differences in performance were identified during testing.<br>- NIST AI bias guidance has been incorporated.<br>- Responsible AI testing and red-teaming have been undertaken.<br>- Ongoing bias assessment remains part of the evaluation framework<br><br>Members emphasised the need for continuous monitoring as AI systems evolve over time. |
| | Concerns were raised regarding:<br>• Older people<br>• Vulnerable users<br>• Individuals with limited digital access<br>• Potential digital exclusion.<br><br>Response: Citizen Assist is an additional reporting mechanism only and traditional reporting routes will remain in place. The Constabulary remains committed to serving all communities through multiple access channels. |

| | |
|---|---|
| | The panel highlighted accessibility and public confidence as key issues for continued monitoring. |
| **6.** | **Scrutiny of live examples** |
| | Sgt Diggle explained that he could present example cases on any criteria. The Panel discussed some criteria they would like to see and the feedback on key findings is as follows:<br><br>Members were shown:<br>- A bicycle theft report<br>- A shoplifting case resulting in a positive criminal justice outcome<br><br>The examples demonstrated that:<br>- AI is used only at the reporting stage<br>- Human review remains central throughout the investigative process<br>- Full transcripts are retained<br>- Existing investigative pathways remain unchanged after submission<br><br>The panel expressed reassurance that AI usage remains constrained, auditable and subject to significant human oversight. |
| **7.** | **Summary & Actions**<br>Panel members were broadly reassured by the controls and governance in place and welcomed the openness shown by the Constabulary in presenting a live AI deployment for external scrutiny. Value was found in discussing transparency, bias, safeguarding, public confidence and evidential integrity.<br><br>Key assurances:<br>- AI is limited to information gathering and triage<br>- Human oversight is maintained throughout<br>- Full transcripts are retained and reviewable<br>- Existing reporting routes remain available |

| | |
|---|---|
| | - Governance, ethics, information security and data protection assessments have been completed<br>- Bias testing and ongoing evaluation are embedded within the pilot<br><br>**Actions:**<br><br>- Consider future analysis on whether user behaviour differs when reporting through AI compared with traditional reporting routes. **Constabulary action.**<br>- Provide panel members with a publishable version of the Citizen Assist evaluation and benefits realisation work when available. **MP to liaise with Constabulary re availability and suitability of this.**<br>- Consider mechanisms for capturing public satisfaction and user feedback from Citizen Assist users. **SN confirmed that this is something IT are working on. Update to be provided when available.**<br>- Bring evaluation findings and benefits realisation information back to the panel at a future meeting. |
| **8.** | **AOB**<br>Next meeting: to be confirmed. |