To whom it may concern,

Please find below my comments regarding TPS' consultation on the use of AI technology.

## Preliminary Comments

It is relevant for TPS to come up with a validation procedure such as this one given that: Machine Learning is still an active field of research, meaning that there is still much that we could understand better; AI has demonstrated its performance more empirically rather than theoretically; and, as such, it does not necessarily offer the same degree of control over the result that it computes, which is due in part to the fact that they heavily depend on the data used as part of their training. Finally, police activities are such that it may be challenging or impossible to know a posteriori with absolute certainty whether a decision or task was properly made (e.g. latent print identification, for which a match having been found or an individual being prosecuted does not demonstrate that the task was performed correctly).

All of this, combined with the duties and activities of police forces, makes it necessary for law enforcement to consider which AI algorithms / systems are appropriate for use, how they should be used, and to which validation process(es) they should be subjected. In that regard, I was pleasantly surprised to hear of your effort to establish a validation process for such algorithms, and of your intent to inform and involve the general public.

## Remarks

Here's a subset of general issues/challenges which come into play when considering how to approve the use of ML/AI algorithms in high-risk settings.

**Which algorithms should be subject to approval?** It is important to specifically define which algorithms / systems will be subject to validation. On what basis is an algorithm selected for review? The fact that it is reported as a ML algorithm by the developer? If so, it can be circumvented. Is it determined by TPS, and if so, on what basis? Will TPS require access to the source code? If an algorithm does not fall within the general field of Machine Learning but shares a subset of its characteristics, including, for instance, dependence on a data set for training / calibration / determining its inner parameters, should it still be subject to approval? This entire conversation is sparked by the fact that automated systems can perform a task and potentially make or assist in a decision which can be detrimental to an individual.

**ML and forensic science**. ML researchers and developers are generally not knowledgeable of the requirements for the admissibility of forensic evidence and the practices across different forensic disciplines, which leads to many suggested "AI systems of security" to circumvent or otherwise ignore the possible legal implications of using such systems. In truth, in applications where AI systems are used where a mistake could have legal repercussions on an individual, there should be legal responsibilities for both the developer and the user depending on the risk of the malfunction. It is therefore necessary for all parties to anticipate these issues and establish best practices.

**Guidelines for the use of AI algorithms**. Given the above remark, it is relevant to consider how AI algorithms used in high-risk settings should be used. It begs the question as to whether the

developer of the system, the user(s) of the system, or both, jointly, design a user manual and/or best practices for the use of such systems (e.g. when making a forensic analysis, should the expert use the algorithm after making their own decision? What if this is not possible?).

**ML approach and data used**. Developers of AI systems often favor short-term approaches and feasibility over quality when designing algorithms intended for use in high-risk settings:
When the algorithm is designed to perform a task already performed by a human expert, insufficient effort is made by the developer to ensure that the algorithm meets the requirements expected of human experts (e.g. AFIS vs latent print examiners, with discrepancies documented as part of the 2009 PCAST report and echoed in AAFS 2017);
Alternatively, notably in recent times, software developers offer algorithms for which its performance can hardly be judged objectively (e.g. determining the sentence of an offender, or "facial reconstructions" from 3D models of a complete cranium).
There are also many documented examples of systems (some even put into production) which make use of poor quality training data, therefore making them subject to several types of bias. In some cases, the means / provenance of the data used is undocumented, unethical, or borderline illegal (e.g. Clearview AI scandal).

**Software updates**. An algorithm or system is not set in stone, and instead evolves. This has been made apparent to the general public in recent years as high-profile websites and software developers continuously improve upon their implementation and deploy it to their users. This begs the question as to which process should be undertaken by the user (TPS) when a software update is released. Is it accepted by default? Does it depend on the nature of the software update? Should the update require a complete re-approval, or an expedited procedure? Perhaps all that is necessary is for the updated software to pass a number of automated software tests (for instance end-to-end tests). If so, what should be the nature of these tests, who is responsible for defining the specifications of these tests, and for developing them?

## Suggestions

Given the above challenges, here are some suggestions which should be further discussed and analyzed based on their efficacy and their feasibility.

- Risk categorization should be done at least in part based on whether it may have legal repercussions on an individual.
- Approval process should be reviewed on the basis of its effectiveness (based on the above challenges) and fairness (AI system provides equal access / chances to have an algorithm approved, e.g. associated fees - if there are any - should be kept to a minimum).
- Collaboration with other law enforcement services should be undertaken or at least encouraged on the topics of the approval of ML algorithms and the establishment of user manuals / guidelines.
- Researchers in ML should be involved in the approval process and its review since this is an active field of research.
- Researchers in Forensic Science should be involved in the approval process of algorithms deemed as high risk, as they can provide an insight into existing practices in forensic disciplines.