

# Equality impact assessment: Template for forces

## Live Facial Recognition

### How to complete an equality impact assessment

The equality impact assessment (EIA) will help you to evidence your public sector equality duty (PSED) compliance. It is a live document that outlines the way equality has been considered throughout the life of the policy, process or other project (hereafter referred to as 'product').

If you need help, see the accompanying guidance or contact your diversity, equality and inclusion team.

### Version control

Version number	Date	Author	Comments
1	30/04/2023	Scott Ashby	
2	09/08/2024	Jennifer Housego	
3	27/05/2025	D/Supt Jennings	
4	19/02/2026	D/Supt Jennings	

### Overview

<b>Name of product under development or review</b>	Live Facial Recognition (LFR)
<b>Description of the product</b>	Live Facial Recognition software takes images obtained from CCTV cameras, takes the biometric data from facial images obtained, and cross refers them to a pre-loaded watchlist of images / biometric data,

	highlighting if it is believed that there is a match between the two.
<b>EIA start date</b>	09/08/2024
<b>EIA author and role</b>	Stephen Jennings – Authorising Officer
<b>EIA owner and role</b>	Stephen Jennings – Authorising Officer
<b>Date of last review</b>	19 <sup>th</sup> February 2026
<b>Date of next review</b>	31 <sup>st</sup> October 2026

## Key product development dates

To evidence PSED compliance, equality and inclusion must be considered throughout the life cycle of a product. You can do this by integrating EIA reviews into key product development stages. List your product's key stages and the dates you will review your EIA here.

<b>Key stage</b>	<b>Date</b>
Procurement phase and pre deployment testing	01/01/2024 to 31/07/2024
Post testing and Initial Deployment	30/09/2024 to 31/10/2024
Cambridge Testing – Controlled sensitivity and specificity testing	01/12/2024 to 31/03/2024 report published 13/03/2026
ISO 42001 Accreditation	17/03/2025
National Physical Laboratory (NPL) testing	21/05/2025 report published 13/03/2026
Professor Fussey – Equitability Assessment	Date to be agreed

## Research and evidence

To undertake effective equality assessment that meets PSED compliance standards, you must work from an evidence base. Use this section to list the research you will use to understand the product's potential or actual equality impacts (for example, surveys, customer feedback, protected characteristic data, academic research).

1. NIST algorithm study results for Corsight algorithm July 2022 (Provided by supplier during procurement in support of PSED evidence)
2. Department of Homeland Security (US) study results for Corsight algorithm July 2022 (Provided by supplier during procurement in support of PSED evidence)
3. Internal Assessment of a Live Facial Recognition Deployment following a false positive alert which occurred on 19<sup>th</sup> October 2024.
4. Cambridge University Study - Controlled sensitivity and specificity testing.
5. ISO 42001 accreditation.
6. Professor Peter Fussey – independent academic review of the LFR evaluation in Essex.
7. Essex University Criminology Dept - Project: 'Understanding Public Perception and the Policing Impact of Live Facial Recognition in Essex'
8. The BSI Algorithm Auditing & Dataset Testing (AA&DT), an independent and impartial assessment of AI algorithms, models and datasets. Tests include Bias in AI systems and AI aided decision making (ISO 24027), Robustness for Neural Networks ISO (ISO 24029-1), Assessment of machine learning Part 1 (ISO 4213), Information Technology and Medical image-based modelling for 3D printing – Segmentation (ISO 3532-2).
9. National Physical Laboratory (NPL) – Accuracy and Equitability Evaluation of Corsight Apollo 4 Live Facial Recognition.

## Consultation and Engagement record

Stakeholder consultation will give you a better understanding of your product's impacts and is crucial to satisfying the PSED requirements. Use this section to record the engagement you have undertaken, summarise the feedback received and note subsequent actions. This section can also be used to record nil returns.

Name of group or organisation	Date of contact	Date reply received	Feedback	Action taken or reason why no action was taken
Essex Police Independent Data Ethics Committee	27/07/2023	27/07/2023	<ul style="list-style-type: none"> <li>- Committee broadly supportive of EP exploring development of LFR capability.</li> <li>- Advises two ethical issues; 1) clear about intended use cases.</li> <li>- 2) importance or being open with the public and engage.</li> </ul>	The force developed detailed policies and procedures to be clear on operational use of LFR. An engagement strategy and media strategy was created.
Strategic Independent Advisory Group	10/10/2023	10/10/2023	<p>The key questions posed from Essex Police to the group;</p> <ol style="list-style-type: none"> <li>1) Would the introduction of LFR take away from local resourcing?</li> <li>2) Would LFR be utilised as part of OP Union?</li> <li>3) Would people wearing masks affect facial recognition software?</li> <li>4) Where does the data come from for the software?</li> </ol>	Feedback considered and actioned during pilot
Essex Police Facial Recognition Ethics Review	22/12/2023	22/12/2023	Outlined the success of the pilot including public support. Importance of continued public engagement.	Points considered during the procurement period and initial deployments.
Essex Police Facial Recognition Ethics Review	16/07/2024	16/07/2024	Fully supportive that the use of LFR is in the public interest. Making sure the staff are aware of policies	All documents published both internally and externally.

			and procedures. Deployment plans were both necessary and proportionate.	
Essex Police Independent Data Ethics Committee	03/12/2024	03/12/2024	Minutes are available	The committee felt it was in the public interest to utilise LFR in Essex. Actions and feedback have assisted shape policy, procedure, and operational deployments.
Essex Work Experience Programme	21/08/2024	21/08/2024	Input provided to thirty secondary school aged children	No negative feedback. Awareness Campaign
CSAS <sup>1</sup> Network Day	02/10/2024	02/10/2024	Attended by partner agencies both public, private, and local councils. (50 present)	No negative feedback. Awareness Campaign
Independent Data Ethics Committee	20/11/2025	20/11/2025	Minutes will be available	The committee recommended the following.  <ul style="list-style-type: none"> <li>- That the Police develop FAQs and a method to simply explain LFR to the public.</li> <li>- Periodically analyse positive interventions.</li> </ul>

---

<sup>1</sup> Essex Police Community Safety Accreditation Scheme (CSAS)

				<ul style="list-style-type: none"> <li>- Provide a simple method of explaining LFR to the public.</li> </ul>
Independent Data Ethics Committee	02/12/2025	20/11/2025	Minutes will be available	<p>The committee recommended the following.</p> <ul style="list-style-type: none"> <li>- Balancing the results of both the Cambridge and NPL testing.</li> <li>- Ensure officers are appropriately trained.</li> <li>- That deployment consideration and testing should be catered for the community of Essex.</li> <li>- That staff should be provided with training about the software to better advise the public.</li> </ul>

## General considerations

*Use this section to note any general diversity, equality and inclusion activities or considerations that are relevant to the product. Please only document general considerations here. Protected characteristic analysis can be completed in the section below.*

The purpose of the Equality Impact Assessment (EIA) is to promote fairness and inclusion to staff and the public considering all protected characteristics. This will help identify and mitigate any potential impacts on specific groups and find ways to reduce or eliminate them. This encourages evidence-based decisions considering the needs and experiences of diverse groups in line with legislation (Public Sector Equality Duty & Equality Act 2010). Leading to transparency and accountability.

## Benefits of LFR

Essex Police have been deploying LFR capability in an operational setting since August 2024. The results of those deployments have shown that LFR is a valuable police tactic to locate people who are either outstanding for crime, or at risk of serious harm, or have protective orders issued by the courts.

Before any deployment of equipment, a watchlist will be created. The watchlist is bespoke for every operation / deployment and the rationale for the make-up of the watchlist must be intelligence led, justified, proportionate and necessary, with the nature of the watchlist recorded prior to each deployment.

The criteria for construction of LFR watchlists must be approved by the Authorising Officer (the 'AO') and be specific to an operation or to a defined policing objective. Watchlists, and any images for inclusion on a watchlist, must also be limited to the categories of image articulated in force policy documents which are images of people who are:

- a) wanted by the courts; and/or
- b) suspected of having committed an offence, or where there are reasonable grounds to suspect that the individual depicted is about to commit an offence or where there are reasonable grounds to suspect an individual depicted to be committing an offence; and/or
- c) subject to bail conditions, court order or other restriction that would be breached if they were at the location at the time of the deployment; and/or
- d) missing persons deemed increased risk; and/or
- e) presenting a risk of harm to themselves or others.

CCTV footage obtained as part of the LFR deployment may identify or indicate a person's race or religion. However, the footage is only used by LFR to obtain biometric data to compare against the watch list and is deleted within 31 days.

The deployments and results to date (version of this report) are as follows.

- 76 deployments
- 2,586,230 faces seen
- 350 Positive Alerts
- 4 False Positive Alert
- 120 Arrests
- 5,125 Public Engagements

The data supported the threshold set by the 'Authorising Officer' of 55 which has remained throughout all deployments. With over two and a half million faces seen, over three hundred fifty positive alerts and four false positives. The balance and importance of the threshold setting is to try, where possible, to avoid any 'false positive alerts' (where the software alerts to a person being on the watchlist but is not the correct person) but also trying not to miss those specific individuals who may be sought by law enforcement. This balance is important to ensure purpose limitation and data minimisation of sensitive biometric processing.

The document refers to 'Threshold Settings' throughout. The terminology means – 'A configured confidence score between 1 and 100 at which the facial recognition system

presents a potential match to a human operator for review. This setting plays a critical role in balancing operational effectiveness with fairness by ensuring that only higher-confidence matches are flagged. Setting the threshold too low can increase the risk of false positives, while setting it too high may result in genuine matches being missed (false negatives). Identifying the right threshold therefore involves carefully balancing these risks to support both accuracy and equitable outcomes.

The Cambridge and NPL studies have been completed with their findings now published. Both evaluations indicated that the algorithm demonstrated strong overall performance, though it showed higher accuracy in identifying Black males compared to White males. Following evaluations by NPL, Cambridge determined that the algorithm deployed at a threshold of .55 maintains equitable performance while achieving a false positive rate significantly below the national and legal requirement. Essex Police have also engaged with the Chief Scientific Advisor to provide expert interpretation and contextual insight, as set out below.

*“Professor Paul Taylor, in his capacity as Police Chief Scientific Adviser, provided advice on interpretation of NPL's results, in respect of specific questions we had around thresholds set at below 50, above 55 and 51. He advised that the analysis suggested we should not use a threshold below 50 unless there was exceptional operational need to maximise identifications. Additionally, the results did not support adopting a threshold above 55, because of the risk of not identifying offenders who are a threat to the public. The threshold could be lowered to 51, but [according to NPL's analysis] the gain of 1.7% in identifications [which] is not necessarily worth increasing the proportion of false alarms by .23%.”*

This extensive testing of the algorithm reflects Essex Polices determination to ensure an enduring assessment of the impact of this technology on our communities. If, through ongoing testing, we determine that increased accuracy and confidence can be achieved through amendments in threshold, we will keep these variables under constant review and amend, audit and log where appropriate.

Work has been undertaken to develop an evaluation framework for LFR which will be subject to independent academic review, and the force is committed to ongoing testing of the software. Essex Police have also sought advice from their own independent Data and Digital Ethics Committee in relation to their use of LFR in general and its deployment on more than one occasion following the pilot, prior to deployment and post initial deployment.

It is important to note that the algorithm is unique to the software being used. Therefore, any threshold setting is bespoke to the specific software and determined following the analysis, which is reviewed regularly. This includes quarterly analysis of positive interventions looking at crime outcomes and breakdown of specialist characteristics such as age, gender, and ethnicity.

## Impact assessment and actions

Apply learning from research, consultation and project knowledge to consider equality considerations relevant to your project. This should include any potential or actual impacts

(positive or negative), as well as how the project will uphold the three PSED aims for each of the **protected characteristics**.

1. Eliminate discrimination, harassment and victimization.
2. Advance equal opportunity.
3. Foster good relations between people of different characteristics.

If you have identified a negative impact, note what mitigating action will be taken to reduce or eliminate that impact. If no mitigating action can be taken, please explain why. The issues or impacts identified may change, or new factors may emerge, as your project develops. Use the EIA to document these and how your project has evolved to accommodate equality considerations.

Don't forget to consider intersectionality. This refers to when characteristics overlap to shape experiences of inequality and discrimination. For further information, please refer to the 'Addressing Intersectionality within Policing' report.

Duplicate the boxes below if required. If you need further information about a protected characteristic, open the webpage linked to each subtitle.

## All characteristics

### Details of positive and/or adverse impact or other issue

Whilst LFR deployments will not target persons owing to any protected characteristics or equality groups, there are issues that may arise because of the deployment that can have an impact. If a group is not listed below, it is because there is no anticipated differential impact on that group:

**Age** – Facial images uploaded will be sourced whether from existing Essex Police records or from family and friends of persons reported as missing. The reference image database, consequently, may have images of subjects that were taken a number of years ago.

The age of criminal responsibility in the UK is 10 years old. Image capture via Custody Imaging on which Facial Recognition Technology (FRT) is reliant, is dependent on the age, date, and time at which the custody image was taken. In addition, the European Union’s Agency for Fundamental Rights ‘Facial Recognition Technology Fundamental Rights Considerations in the Context of Law Enforcement Report 2019’ highlights that as a child grows and time passes, the accuracy of a biometric match can diminish. The risk of a failure to match increases when facial images recorded at young ages are compared more than five years after collection. The report further indicates that the accuracy of FRT is in general significantly lower for children younger than 13 years old. They associate this to “rapid growth and change in facial appearance.” Essex Police recognise the increased sensitivity of biometric processing of children and any authorised such deployments will receive detailed focus from the AO to ensure necessity and proportionality.

**Disability** - People can undergo facial change for assorted reasons. They may suffer facial disfigurements through trauma or a medical intervention, or their face may have reconstructive surgery which would result in a significant change to their facial features. Genetic conditions such as neurofibromatosis also cause progressive facial change. Consequently, the images that Essex Police hold may not accurately reflect their present facial appearance. Should such circumstances present themselves to the AO – a flag may be placed against the person of interest to notify operators and responders of potential for reduced accuracy.

**Gender reassignment** – The FRT probe image is based on the mapping of key facial indicators when comparing a reference image database image for an individual. Therefore, the functionality, accuracy, and performance of FRT may be less effective if changes to facial appearance have occurred between the time the reference image used in the database image was taken, and the time a comparison is made.

This may impact persons who are transitioning from one gender to another if gender presentation differs from the time the comparator image was taken. It may also affect trans, non-binary and gender- fluid people who adopt to flex between gender presentations. Reports suggest that facial contouring using cosmetic make-up application may impact on FRT system’s performance.

**Racial Groups** – Essex is a diverse multi-cultural county which incorporates both rural and metropolitan areas. It is therefore important to ensure that the technology is not seen to cause division between persons of different race / ethnicity.

FRT is based on the mapping of key facial indicators. They are also dependant on the ability of the algorithm to determine the key facial indicators within an image. This can be impacted by environmental factors such as ambient light and shadows factors. This may also be impacted by the depth of skin pigmentation, and the use of contouring make up.

To date ethnicity biases have received considerable attention, particularly from academics and government bodies. Relevant studies include Klare et al (2012), NIST (2018) and Buolamwini and Gebru (2018). The findings from Buolamwini and Gebru’s study were widely reported, as they found algorithms were particularly biased in terms of gender and ethnicity; performance was best for men and white individuals, and poor for women and black individuals.

**Religion or Belief** - The wearing of religious headwear or coverings and the growing of facial hair may have an impact on the effectiveness of FRT. In addition, certain cultures or sexes within a religion i.e. Amish, refuse to allow themselves to be photographed. Sensitivity therefore needs to be taken with cross-community dialogue to ensure the deployment is both necessary and proportionate.

Respect for diversity awareness training is embedded Essex Police Culture.

**Sex** - Social observation indicates women change their appearance more frequently and significantly than men which may impact the performance of LFR. Reports suggest that facial contouring using cosmetic make-up application may impact on the LFR system’s performance.

**Mitigating action for any adverse impact or rationale for no further action**

**Equitability**

**Procurement Phase and Pre-deployment Testing**

The selection of an LFR provider by Essex Police had a strong focus on the need to satisfy the Public Sector Equality Duty (PSED), and in this context the chosen provider of the LFR software was asked to provide details of any independent bias testing that had been conducted. Currently there is no nationally agreed standard for LFR, and no nationally approved or mandatory testing is available until very recently. It is therefore up to individual police forces to make reasonable endeavours to obtain details of any independent testing that has been conducted in relation to their chosen software from the supplier and to ensure that the software is as accurate and free from bias where possible.

Essex Police values commitment to ongoing testing from its LFR supplier. This reflects its determination to reflect the ongoing and enduring responsibility to comply with PSED. Our commitment to distil learning from such testing into its deployment strengthens our approach.

In addition to ensuring that independent bias testing has taken place, and that the results of the testing are acceptable to the force in terms meeting the PSED. Essex Police have also developed robust operational policy and procedures to ensure that the criteria for the deployment of LFR and the way it is used by officers helps the force further reinforce their compliance with the PSED.

Essex Police have contracted with Digital Barriers for LFR who provide LFR software from Corsight. Corsight have provided information in relation to the current and complete independent testing of their LFR software and ongoing or future plans for such testing.

Corsight has advised that their algorithms have been vigorously tested independently for both accuracy and bias by the National Institute of Standards and Technology (NIST) and the Department of Homeland Security (DHS) in the US and have provided details of the results of testing in 2022. In both tests, Corsight's algorithms were found to have one of the lowest bias differentials available of any algorithm tested. Bias differential is a calculation that distinguishes False-Match Rates (FMR) between varying demographic groups (gender, skin tone) to distinguish if there is more, or less, likelihood of a false-match when comparing two images of different demographics. The aim of all algorithms is to have a small as possible FMR difference between all demographic categories to ensure there is no more likelihood of a false match between one demographic or another.

To measure this, NIST calculates what is known as the False Match Rate (FMR) — how often the system incorrectly says two different people are the same. They assess this across four key demographic groups: White males, Black males, White females, and Black females. A fair system should make errors at about the same rate for all of them.

To allow for consistent comparisons across vendors, NIST sets a common baseline false match rate of 0.0010 (1 in 1,000), based on the group that tends to perform best — usually White males. Then, using that common threshold, NIST measures how each demographic group performs with every algorithm.

For Corsight's algorithm ("Corsight\_003"), the results were:

- White Male: 0.0010 (1 in 1,000)
- Black Male: 0.0007 (1 in ~1,429)
- White Female: 0.0005 (1 in 2,000)
- Black Female: 0.0011 (1 in ~909)

When analysing the results, a bias differential can be calculated based on the difference between the highest and lowest FMRs to understand how evenly the algorithm performs across demographics.

- Highest FMR: 0.0011 (Black Female)
- Lowest FMR: 0.0005 (White Female)
- Bias differential:  $0.0011 - 0.0005 = 0.0006$

This bias differential of 0.0006 means that the largest difference in error rates between any two demographic groups was just 6 in 10,000 comparisons, or roughly 1 in 1,667.

The closer this value is to zero, the more evenly the system treats people from different demographic groups.

At the time of testing, this was the lowest bias differential observed across all submitted algorithms — demonstrating that Corsight’s technology was among the most consistent and fair-performing systems tested.

Corsight also participated in the challenging and highly competitive U.S. Department of Homeland Security (DHS) Biometric Technology Rally in September 2022. For the purpose of anonymised evaluation and comparison, Corsight was assigned the pseudonym “Tioga,” as per the event protocol. The official results are publicly available on the DHS website.

This focussed on testing biometric algorithms in realistic, high-traffic conditions based on varying qualities of image acquisition systems. The Corsight algorithm was top ranked for accuracy and for being unbiased toward gender or race based on their test conditions. DHS testing is based on system aliases ... It achieved a 99.8% overall Matching True Identification Rate (TIR) and over 99.3% TIR for all skin tone groups across all acquisition systems, and under some acquisition systems had 100% TIR, demonstrating Corsight’s capability to perform equally across all demographics.

Intertek NTA conducted a pre-deployment IT health and internal security check.

Essex Police completed pre-acceptance testing (MVP) which took place on 23<sup>rd</sup> / 24<sup>th</sup> July 2024 as part of the pre-acceptance phase. The criteria of testing were as follows:

- Face Detection Accuracy
- Face Recognition Accuracy
- Performance Under Load
- Alert Generation
- Light / Weather Conditions
- Audit Testing
- System Performance
- Functionality Test
- Hardware / Software Installation
- Configuration Verification / Security Testing Access Control

The tests included police officers from different age ranges, sex, and ethnicity. These localised results indicated that the software / algorithm was safe to use at a threshold setting of 0.45 where there were no ‘false positives’ recorded.

The PSED duty requirement is continuing responsibility and Essex Police acknowledged that further and ongoing testing would further assist. Therefore, the threshold was set at 55 for all deployments. This has remained the same until present day. This meant that at this setting the software will identify those who are sought during a deployment but limit the possibility of falsely identifying someone who is not.

### **Post testing and Initial Deployment**

On 19<sup>th</sup> October 2024, LFR was deployed in Southend on Sea city centre. During this deployment there were 1322 subjects on the approved watchlist. During the deployment

39,401 faces were scanned. There were five positive alerts leading to three arrests. One of the positive alerts led to an approach to a member of public who was not the person on the watchlist. This was categorised as a false positive.

This is the first time that a false positive has occurred since inception and use of LFR in the county of Essex.

A full analysis took place to understand why a false positive had occurred. The results showed that the image uploaded to software of the person sought was a 'thumbnail' version and not the original. Therefore, the pixelation was lower and led to the false positive. Working with Corsight, the software has now been re-configured to prevent an image below a certain level being uploaded to the watchlist which will prevent a repeat of false positive.

As part of the review of the false positive, the CCTV footage was analysed with the same watchlist in a controlled environment at the police station but varying the threshold levels. The results of 0.40 threshold test showed a total of 9 individuals were subject of positive alerts, 5 of which matched the positive alerts from the live deployments. The additional 4 were not subjected of alerts during the live deployment which suggested that they were missed due to the threshold set. The images on the watchlist and on the camera footage were visually viewed and appeared to be the same people but not conclusive. This suggested that there were no issues with the software and showed that four individuals on the watchlist were missed at the threshold setting of 55. They would have been alerted at 0.40.

Between 18<sup>th</sup> and 26<sup>th</sup> August 2025 there were three False Positive Incidents as outlined below:

18<sup>th</sup> August – A low quality image was uploaded onto the system and provided a match. The image was only 1.9kb which should have been declined by the system.

23<sup>rd</sup> August – The image was 9kb so of reasonable quality. The system gave a match level of 55.19 confidence which is just over the level set at 55. A visual inspection of the photo and the person stopped does show significant similarities.

26<sup>th</sup> August – The image was 216kb so well in excess of the recommended level of 100kb. The image was side profile of a high-risk missing child and again a visual review suggests the image from COMPACT and the person stopped were very similar in appearance. The system match was at 75.46.

Following the false positive alerts noted above, LFR deployments were paused on 26<sup>th</sup> August to allow for thorough analysis and the implementation of corrective actions.

Corsight have added Camera Groups to allow easier management of camera settings (adding new wide lens to the cameras, for example).

System settings have been updated so that a live image will only trigger facial recognition once the face reaches a resolution of 60 pixels, increased from the previous threshold of 50 pixels. This adjustment enables a more accurate assessment of the image before any alert is generated

Settings have been adjusted so that the minimum accepted face size of enrolment of pictures is 100 pixels. This will filter out any image with a face size smaller than 100 pixels wide.

### **Cambridge Testing – Controlled sensitivity and specificity testing**

Cambridge University have carried out a study examining the accuracy of the Corsight Algorithm to further assist Essex Police ensuring they comply with their PSED requirements.

The study took place in February 2025 using 188 actors of whom sex, age and ethnicity varied to test if accuracy differs by demographic. This was a broad replication of the NPL testing of the NEC software.

This was looking at a) the efficiency of the software in identifying people we knew were present and b) the inefficiency of the software in identifying people it should not have. The design is intended to replicate real-world conditions of an LFR deployment.

The study addresses three research questions.

- a. What is the overall sensitivity and specificity of the Corsight software at different confidence thresholds?
- b. What (if any) is the disparity in sensitivity and specificity rates between (i) age groups, (ii) sexes, and (iii) ethnicity groups
- c. What is the association between occlusions and any disparities identified in b.

The report provided the following results at a threshold setting of 55.

- Correctly identified known to be present males more often than females.
- Correctly identified known-to-be-present individuals of Black ethnicity more often than those of White ethnicity by a rate of about 1 in 10.
- In other words, if 100 Black and 100 White subjects passed the camera, 10 more Black people would be correctly identified than White people.
- Disparity doesn't exist in incorrect identification.
- No disparity in interventions - arrests etc.
- The report recommended that the threshold could be reduced safely still avoiding any false positives.

Following receipt of the report, several mitigating actions were initiated which included the following.

- 1) Engagement with Digital Barriers and Corsight re the bias linked to the True Positive Identification Rate (TPIR). Corsight have analysed and the algorithm and reduced the level of bias linked to the TPIR.
- 2) Consultation with the Data Ethics Committee.

- 3) Awaiting the final report from the National Physical Laboratory (NPL) to compare the results.
- 4) Stakeholder engagement to assist interpretation of the results balancing against operational settings.
- 5) Live deployments were paused to complete the above actions.

Since the results of the Cambridge Study showing some bias linked to the True Positive Identification Rate (TPIR), our response included engagement with our software provider who sought to make improvements to the algorithm considering the findings. The provider has since confirmed that they have made advancements to reduce disparity and are working on further improvements to the algorithm.

### **National Physical Laboratory (NPL) - Accuracy and Equitability Evaluation of Corsight Apollo 4 Live Facial Recognition.**

The NPL study closely replicated the methodology applied to the NEC algorithm in 2022 after the Bridges case. Notably, the footage and watchlists used to evaluate the Corsight algorithm were identical to those utilized in the five 2022 deployments—four in London and one in South Wales. The evaluation was conformant with ISO accreditation (ISO/IEC 19795-1 (1) and ISO/IEC 19795-2 (2)).

The study examines recognition accuracy in two forms – True Positive Identification Rate (TPIR) the total number of times an individual(s) on a watchlist who is known to have passed the LFR system and correctly generate an alert and False Positive Identification Rate (FPIR) the number of individuals who pass the LFR system, but who are not on the watchlist and who incorrectly generate an alert.

The study predominantly focussed on Essex Police operational threshold setting of 55 but helpful provides context at other thresholds for comparison purposes. Equally the study provides results comparing against different watchlist sizes – 180,000, 18,000 and 1,800. It is important to note that Essex Police to date have deployed with on average 2,000 on a watchlist. It is also important to note that the demographic breakdown of the 180,000-watchlist used in the study was roughly one third white individuals, one third individuals of black ethnicity and one third of 'other ethnicity'.

Statistical significance in evaluation is assessed at conventional significance level of 0.05 (5%).

The key headlines are summarised below.

- At a threshold of 55, 60, and 65 the algorithm is “equitable, i.e., no statistically significant bias” in TRIR and FPIR.
- At 55 there is no “statistically significant” demographic variation. (this is with a watchlist of 180,000)
- At a threshold of 55 and watchlist size of 18,000 the software will identify 89% of all persons present and falsely identify 0.017% (1 in 5,700)
- At a threshold of 50 (and lower) demographic variation was deemed “statistically significant”.

- The NPL found statistically significant variation in the TPIR results when there were environmental condition changes. (93% at Piccadilly circus on 28/07/22 compared to 84% at Oxford Circus 07/07/22).
- When the threshold is set at 60 and above the overall TPIR starts to decrease rapidly.
- The report states - *"Whether the impact on a data subject is negligible or major concern is generally specific to the operational use case. For LFR, false alerts, which may direct a police officer to engage with the data subject, are generally of greater impact on the data subject than missed identifications. Consequently, demographic variation would likely be of concern if one demographic group has a FPIR in excess of the norm."*
- The final summary states "Under our equitability criteria, the algorithm is considered equitable at thresholds 65, 60, and 55. At threshold 50 and lower, equitability will depend on watchlist size and other deployment settings."

Essex Police engaged with Professor Paul Taylor – the Chief Scientific Advisor to assist with technical interpretation and context.

Professor Paul Taylor, in his capacity as Police Chief Scientific Adviser, provided advice on interpretation of NPL's results, in respect of specific questions we had around thresholds set at below 50, above 55 and 51. He advised that the analysis suggested we should not use a threshold below 50 unless there was exceptional operational need to maximise identifications. Additionally, the results did not support adopting a threshold above 55, because of the risk of not identifying offenders who are a threat to the public. The threshold could be lowered to 51, but [according to NPL's analysis] the gain of 1.7% in identifications [which] is not necessarily worth increasing the proportion of false alarms by .23%.

### **ISO 42001 and Biometric System Testing (BSI) Accreditation**

The ISO/IEC 42001 standard, set to guide AI practices, helps ensure compliance with principles of AI accountability, transparency, and governance. Development of AI relies upon interaction between ethical development, lawfully acquired databases and quality managed systems geared to digest and implement emerging regulation. Essex Police is aware that the regulatory, legal and ethical landscape for AI is fast changing. Ensuring any supplier of AI implements the disciplines of ISO 42001 is of significant importance to Essex Police as part of our reassurance to communities. The National Police Chiefs Council Approved Professional Practice document interpreted regulatory guidance following Bridges. Essex Police is aware that much of that guidance has been distilled into Corsight's development practices. On 17<sup>th</sup> March 2025 the Corsight software achieved the ISO 42001 accreditation. The algorithm has passed the ISO/IEC 19795-1 in 2025.

### **Professor Pete Fussey (Southampton University) & Dr Daragh Murray (University of London)**

Professor Fussey will be conducting independent academic review of the LFR evaluation in Essex. The work will focus on issues of ethics, human rights compliance, and governance and is designed to complement the ongoing technical evaluation. The main focus will be on:

- 1) Evaluation – proposed plans for LFR technology use.
- 2) Deployment & Operations – advice and guidance on operational policy, practice and deployments.

- 3) Outreach & Engagement - to provide advice and guidance on the approach to engagement with the local community to increase trust and confidence as well as transparency.

**Essex University Criminology Department – Project: ‘Understanding Public Perception and the Policing Impact of Live Facial Recognition in Essex’**

This research evaluated the impact of the LFR technology on policing in Essex through location-based law enforcement deployments.

The project will look at analysing three main areas of interest:

- 1) The public
- 2) Local businesses
- 3) Social media

The study used a mixed-method approach, combining a main survey for the general public and local businesses, follow-up interviews, and a social media analysis of public comments and news on popular platforms (e.g., Facebook, Twitter). The fieldwork was conducted during several LFR deployments in Essex within the months of July and August 2025, and at other locations where LFR had been deployed in the past, gathering 115 survey responses from the public, 20 from local businesses, and 11 follow-up interviews. Furthermore, approximately 1,060 comments from Facebook, Twitter, Reddit, and local news outlets on reputable websites (e.g., The Guardian, Your Southend, Echo Essex) were also analysed to understand the informal online discourse and debates of LFR.

The results in summary provided the following.

- Trust & Legitimacy – supported ‘trust’ is strongest driver.
- Awareness & Understanding – partially supported, ‘awareness’ linked to support.
- Privacy Concerns – strongly supported, ‘privacy’ is the main barrier to support.
- Transparency – supported but interview results show mixed views.
- Local Businesses – partially supported, ‘trust’ matters but ‘customers fears’ reduce support.
- Online Discourse – supported, online discourse more critical and rights-oriented than field responses.

In line with the recommendations from the Cambridge testing, a quarterly review of positive alerts, interventions and outcomes will now take place analysing the equitability of performance across special characteristics of individuals (age, gender, ethnicity) who are subject to alert and intervention.

**Code of Ethics**

The Code of Ethics sets and defines the exemplary standards of behaviour for everyone who works in policing. The Code of Ethics is about self-awareness, ensuring that everyone in policing feels able to always do the right thing and is confident to challenge colleagues irrespective of their rank, role or position.

## Age

<b>Details of positive and/or adverse impact or other issue</b>
<p>Facial images uploaded will be sourced from existing Essex Police records or from family and friends of persons reported as missing. The reference image database, consequently, may have images of subjects that were taken several years ago.</p> <p>The age of criminal responsibility in the UK is 10 years old. Image capture via Custody Imaging on which FRT is reliant, is dependent on the age, date and time at which the custody image was taken. In addition, the European Union's Agency for Fundamental Rights 'Facial Recognition Technology Fundamental Rights Considerations in the Context of Law Enforcement Report 2019' highlights that as a child grows and time passes, the accuracy of a biometric match can diminish. The risk of a failure to match increases when facial images recorded at a young age are compared more than five years after they were collected. The report further indicates that the accuracy of FRT is in general significantly lower for children younger than 13 years old. They associate this to "rapid growth and change in facial appearance". Therefore, during live deployments children under the age of 13 are generally not on the watchlist unless there was significant threat, risk or harm. For example, high risk missing child.</p>
<b>Mitigating action for any adverse impact or rationale for no further action</b>
<p>Any watchlist created will be done so as close to the deployment as possible, therefore hoping to ensure the most accurate and up to date images of persons being added are uploaded. If the LFR software highlights a potential match, police personnel will check the images highlighted for accuracy.</p> <p>At a threshold of 55, the testing shows the false positive identification rate was equitable between all age groups tested.</p>

## Disability or neurodiversity

<b>Details of positive and/or adverse impact or other issue</b>
<p><b>Disability</b> - People can undergo facial change for several reasons. They may suffer facial disfigurements through trauma or a medical intervention, or their face may have reconstructive surgery which would result in a significant change to their facial features. Genetic conditions such as neurofibromatosis also cause progressive facial change. Consequently, the images that Essex Police hold may not accurately reflect their present facial appearance.</p>
<b>Mitigating action for any adverse impact or rationale for no further action</b>
<p>Should a person's face have changed for any of the reasons above (or an additional / different reason) between their image being taken / provided and an LFR deployment, it would ultimately mean that the LFR software would be less likely to highlight a match between the image taken during deployment and that on the watchlist. Should a match be</p>

highlighted, this is then checked by police personnel for accuracy before additional activity is considered. The reality is that if we were trying to locate someone for safeguarding purposes, wanted persons etc, the software may not highlight a match due to a facial change. There are no actions that can be undertaken to stop this from happening.

## **Gender reassignment**

### **Details of positive and/or adverse impact or other issue**

The Facial Recognition Technology (FRT) probe image is based on the mapping of key facial indicators, when compared to a reference picture in the watchlist database for an individual. Therefore, the functionality, accuracy, and performance of FRT may be less effective if changes to facial appearance have occurred between the time the reference image database image was taken, and the time a comparison is made.

This may impact persons who are transitioning from one gender to another if gender presentation differs from the time the comparator image was taken. It may also affect trans, non-binary and gender- fluid people who adopt to flex between gender presentations. Reports suggest that facial contouring using cosmetic make-up application may impact on FRT system's performance.

### **Mitigating action for any adverse impact or rationale for no further action**

Should a person's face have changed for any of the reasons above (or an additional / different reason) between their image being taken / provided and an LFR deployment, it would ultimately mean that the LFR software would be less likely to highlight a match between the image taken during deployment and that on the watchlist. Should a match be highlighted, this is the checked by police personnel for accuracy. The reality is that if we were trying to locate someone for safeguarding purposes, wanted persons etc, the software may not highlight a match due to a facial change. There are no actions that can be undertaken to stop this from happening.

## **Marriage and civil partnership**

### **Details of positive and/or adverse impact or other issue**

There is no anticipated differential impact based on marriage or civil partnership.

### **Mitigating action for any adverse impact or rationale for no further action**

N/A

## Pregnancy and maternity

<b>Details of positive and/or adverse impact or other issue</b>
There is no anticipated differential impact based on pregnancy or maternity.
<b>Mitigating action for any adverse impact or rationale for no further action</b>
N/A

## Ethnicity

<b>Details of positive and/or adverse impact or other issue</b>
<p>Essex is a diverse multi-cultural county which incorporates both rural and metropolitan areas. It is therefore important to ensure that the technology is not seen to cause division between persons of a different race/ethnicity.</p> <p>FRT is based on the mapping of key facial indicators. It is also dependant on the ability of the algorithm to determine the key facial indicators within an image. This can be impacted by environmental factors such as ambient light and shadows. This may also be impacted by the depth of skin pigmentation, and the use of contouring make up.</p> <p>To date ethnicity biases have received considerable attention, particularly from academics and government bodies. Relevant studies include Klare et al (2012), NIST (2018) and Buolamwini and Gebru (2018). The findings from Buolamwini and Gebru's study were widely reported, as they found algorithms were particularly biased in terms of gender and ethnicity: performance was best for men and white individuals, and poor for women and black individuals.</p>
<b>Mitigating action for any adverse impact or rationale for no further action</b>
<p>The algorithm has undergone significant testing both before, during initial deployment and more recently by both Cambridge University and the NPL. The decision has been made taking into consideration both studies that at a threshold setting of 55 will continue to be utilised. At this level, the algorithm is both equitable and the risk of a false positive is far less than the national requirement.</p> <p>If the LFR software highlights a potential match, police personnel will check the images highlighted for accuracy before additional activity is considered.</p>

## Religion or belief

<b>Details of positive and/or adverse impact or other issue</b>
The wearing of religious headwear or coverings and the growing of facial hair may have an impact on the effectiveness of FRT. In addition, certain cultures or sexes within a religion

i.e. Amish, refuse to allow themselves to be photographed. Sensitivity therefore needs to be taken with cross-community dialogue to ensure the deployment is both necessary and proportionate.

Respect for diversity awareness training is embedded Essex Police culture.

**Mitigating action for any adverse impact or rationale for no further action**

Any watchlist created will be done so as close to the deployment as possible, therefore hoping to ensure the most accurate and up to date images of persons are uploaded. If the LFR software highlights a potential match, police personnel will check the images highlighted for accuracy.

Should a person be wearing some form of headwear or covering, it could mean that the LFR software would be less likely to highlight a match between the image taken during deployment and that on the watchlist. The reality is that if we were trying to locate someone for safeguarding purposes, wanted persons etc, the software may not highlight a match. There are no actions that can be undertaken to stop this from happening.

## Sex

**Details of positive and/or adverse impact or other issue**

Social observation indicates women change their appearance more frequently and significantly than men which may impact the performance of LFR. Reports suggest that facial contouring using cosmetic make-up application may impact on the LFR system's performance.

**Mitigating action for any adverse impact or rationale for no further action**

Any watchlist created will be done as close to the deployment as possible, therefore hoping to ensure the most accurate and up to date images of persons are uploaded. If the LFR software highlights a potential match, police personnel will check the images highlighted for accuracy.

Should a person's face / image have changed for any reason between their image being taken / provided and an LFR deployment, it would ultimately mean that the LFR software would be less likely to highlight a match between the image taken during deployment and that on the watchlist. Should a match be highlighted, this is checked by police personnel for accuracy. The reality is that if we were trying to locate someone for safeguarding purposes, wanted persons etc, the software may not highlight a match due to a facial change. There are no actions that can be undertaken to stop this from happening.

## Sexual orientation

**Details of positive and/or adverse impact or other issue**

There is no anticipated differential impact based on sexual orientation.

<b>Mitigating action for any adverse impact or rationale for no further action</b>
N/A

## Socio-economic background

Our socio-economic background is a combination of different factors, such as our income, occupation and social background. A person’s socio-economic background can expose them to inequalities, so it’s important to understand how your project could affect different socio-economic groups. Further information about socio-economic status can be found in the Cabinet Office [Measuring socio-economic background in your workforce](#) guidance.

<b>Details of positive and/or adverse impact or other issue</b>
There is no anticipated differential impact based on socio-economic background.
<b>Mitigating action for any adverse impact or rationale for no further action</b>
N/A

## Other characteristics

Use this section to consider the PSED aims and any impacts your product may have on characteristics that are not protected under the Equality Act 2010 but are still significant to equality and inclusion. For example, your product may have a particular effect on people with caring responsibilities or on people with English as an added language, or you may need to consider the Welsh Language (Wales) Measure 2011 in delivery. Think creatively and invite input from stakeholders.

<b>Details of impact or other issue</b>
There are no additional anticipated differential impacts.
<b>Mitigating action or rationale for no further action</b>
N/A

## Action log

Record EIA actions and monitor action progress in the optional action log.