BIG
BROTHER
WATCH

Big Brother Watch's written evidence for the Justice and Home Affairs Committee's Inquiry into new technologies and the application of the law

September 2021

About Big Brother Watch

Big Brother Watch is a civil liberties and privacy campaigning organisation, fighting for a free future. We're determined to reclaim our privacy and defend freedoms at this time of enormous technological change.

We're a fiercely independent, non-partisan and non-profit group who work to roll back the surveillance state and protect rights in parliament, the media or the courts if we have to. We publish unique investigations and pursue powerful public campaigns. We work relentlessly to inform, amplify and empower the public voice so we can collectively reclaim our privacy, defend our civil liberties and protect freedoms for the future.

Contact

Silkie Carlo

Director

Direct line: 020 7340 6042

Email: silkie.carlo@bigbrotherwatch.org.uk

Acknowledgements: With thanks to Esme Harrington

INTRODUCTION

We welcome the opportunity to submit evidence to the Committee's inquiry into new technologies and the application of the law. In this submission, we wish to share our research on the use of new technologies and the law, particularly on technologies that relate to policing and the criminal justice system. Our submission focuses on three aspects in this regard: live facial recognition, predictive policing, and digital evidence gathering.

Questions

1. Do you know of technologies being used in the application of the law? Where? By whom? For what purpose?
    a. **Live Facial Recognition**
        i. Facial recognition technology measures and matches unique facial characteristics ('biometrics') for the purposes of biometric surveillance or identification.
        ii. There are two types of facial biometric recognition:
            1. Facial matching or 'static' facial recognition: this is the matching of an isolated, still image of an individual against a database. This is used at borders with biometric passports and by police to match images of suspects against images on the Police National Database.
            2. Live facial recognition surveillance: this technology matches faces on live surveillance camera footage against a database (such as the custody image database, or a subsidiary 'watchlist') in real time.
                A. South Wales Police describes the live facial recognition process as follows: The process can be broken down into three very general steps. First, the computer must find the face in the image. It then creates a numeric representation of the face based on the relevant position, size and shape of facial features. Finally, this numeric map of the face in the image is compared to a database of images of identifies faces.
        iii. In the UK, live facial recognition surveillance technology has been deployed by the Metropolitan Police, South Wales Police, Greater Manchester Police, Leicester Police and Humberside Police.

iv. Since 2016, the Metropolitan Police and South Wales Police have deployed this surveillance technology prolifically: at sports matches, concerts, shopping centres and high streets, Notting Hill Carnival, Remembrance Sunday – and even a peaceful demonstration. South Wales Police has received £2m in funding from the Home Office to lead the deployment of automated facial recognition. [1]

v. In 2018, Greater Manchester Police deployed the technology at the Trafford Centre shopping centre for a period of 6 months in 2018 biometrically scanning an estimated 15 million people, before the Surveillance Camera Commissioner intervened. [2]

vi. As of February 2020, the trials had so far cost the Metropolitan Police over £240,000 just in material hardware and software costs, not including the significant costs of teams of uniformed and plainclothes officers in attendance at each deployment.[3] Police have refused to provide the full costs.

vii. The Metropolitan Police announced on 24th January 2020 that it was rolling out the technology operationally across London.[4] However, it has not been used since February 2020 owing to the pandemic.

viii. In July 2021, it was revealed that Hampshire Constabulary, Humberside Police, North Wales Police and South Yorkshire Police have been trialling software developed by company Reveal which provides retrospective facial recognition. This software can analyse publicly-provided or police bodycam pictures and videos to cross-compare with police databases to identify suspects.[5] This significantly increases the scope for privacy intrusion in the course of police encounters.

ix. Collaboration between police and private companies

   1. Several UK police forces have also collaborated with private companies using facial recognition surveillance.

[1] South Wales Police and Crime Commissioner, 'Medium Term Financial Strategy 2017-2021', 28 December 2016 https://pcclivewww.blob.core.windows.net/wordpress-uploads/2016-12-28-Final-Medium-Term-Financial-Strategy.pdf

[2] Working together on automatic facial recognition – Tony Porter, Surveillance Camera Commissioner, 10 October 2018 - https://videosurveillance.blog.gov.uk/2018/10/10/working-together-on-automatic-facial-recognition/

[3] https://www.parliament.uk/business/publications/written-questions-answers-statements/written-question/Lords/2020-02-04/HL1335/

[4] https://www.theguardian.com/technology/2020/jan/24/met-police-begin-using-live-facial-recognition-cameras

[5] https://inews.co.uk/news/technology/uk-police-testing-retrospective-facial-recognition-identify-criminals-1128711.

2. In Sheffield, South Yorkshire Police shared images with Meadowhall Shopping Centre during a secret trial of facial recognition surveillance.[6] Millennium Point conference centre in Birmingham stated in their privacy policy that they used facial recognition "at the request of law enforcement",[7] which they then subsequently denied and removed.
3. Meanwhile, the World Museum in Liverpool initially admitted to trialling the technology "following advice from Merseyside Police and local counter terrorism advisors", which both also then later denied.[8]
4. The Metropolitan Police and British Transport Police shared images with the Kings Cross Estate, which secretly used facial recognition surveillance encompassing one of the country's busiest national and international rail networks, and a large office and retail area.[9]

b. Predictive Policing

i. Geographic Crime Prediction System

1. Geographic crime prediction systems use crime data to create future predictions of where and when certain crimes will occur.
2. One such example is the commercial geographic crime prediction toll created by the company PredPol, which styles itself as 'The Predictive Policing Company'.[10] The eponymous PredPol product feeds crime and location information into a machine-learning algorithm to calculate predictions of times and locations ('hotspots') where specific crimes are most likely to occur. The algorithm is based on an 'earthquake' model of crime that predicts certain crimes result in further 'aftershock' crimes within the same area.[11] The system uses current and historical police crime data to create its predictions: crime type, crime location and crime date and time.[12]

[6] https://www.bbc.co.uk/news/technology-51268093
[7] https://www.itv.com/news/central/2019-08-16/facial-recognition-technology-allegedly-used-at-birmingham-conference-centre/
[8] https://www.liverpoolecho.co.uk/news/liverpool-news/controversial-facial-recognition-used-during-16769707
[9] https://www.ft.com/content/8cbcb3ae-babd-11e9-8a88-aa6628ac896c
[10] https://www.predpol.com/
[11] https://www.predpol.com/
[12] https://www.predpol.com/technology/

3. PredPol was used by Kent Police for 5 years between 2013 and 2018 before it was scrapped, with a superintendent saying it had been 'challenging' to show whether crime was actually reduced as a result.[13] Greater Manchester Police, West Midlands Police, West Yorkshire Police and the Metropolitan Police have also either trialled PredPol or other similar geographic crime prediction systems including their own bespoke systems.[14]

ii. Individual-oriented crime prediction: Durham Constabulary's Harm Assessment Risk Tool (HART)

1. Durham Constabulary has developed its own machine-learning algorithm, the Harm Assessment Risk Tool (HART), which profiles suspects to predict their risk of re-offending in the future, giving them a risk score: high, moderate or low. This AI-generated risk score is used to advise whether to charge a suspect or release them onto a rehabilitation programme, 'Checkpoint'. If individuals who have been assessed by HART as 'moderate' risks successfully complete the 'Checkpoint' rehabilitation programme, they will not receive a criminal conviction. This system therefore has significant consequences for individuals' criminal justice outcomes. The principle of using historic data about an individual to make predictions about their potential future behaviour also brings into question the presumption of innocence and the right to a fair trial.
2. The HART algorithm is based on a random forest model, constructed from 509 separate classification and regression decision trees (CART), which are combined into the forecasting model. HART was built on a dataset using approximately 104,000 custody events over a five year period. It uses 34 different predictor variables to arrive at a forecast, 29 of which focus on the individual's history of criminal behaviour. A further variable is the number of police intelligence reports relating to the individual. The other variables include age, gender and two types of residential postcode.

[13] https://www.bbc.co.uk/news/uk-england-kent-46345717

[14] https://www.ibtimes.co.uk/predictive-policing-predpol-future-crime-509891

3. Big Brother Watch's investigation found that one of the postcode variables fed into the HART system is a commercial marketing data product from the global data broker Experian, known as 'Mosaic'.[15] Mosaic is a socio-geodemographic segmentation tool, consisting of postcode stereotypes created from 850 million pieces of data, including census data, ethnicity, health data, employment, GCSE results, child benefits and income support, family and personal names linked to ethnicity, data scraped from online sources including pregnancy advice websites and much more.[16]
4. This data is used to profile all 50 million adults in the UK[17] into stereotypes based on their postcodes, creating household profiles which, in 2018, included categories such as "Asian Heritage", "Disconnected Youth", "Crowded Kaleidoscope", "Families with Needs" or "Low Income Workers".[18] Experian's profiles attribute 'demographic characteristics' to each stereotype. For example, 'Asian Heritage' individuals were characterised as being part of "extended families" living in "inexpensive, close-packed Victorian terraces", and that "when people do have jobs, they are generally in low paid routine occupations in transport or food service".[19] 'Crowded Kaleidoscope' were described as "multi-cultural" families likely to live in "cramped" and "overcrowded flats", with names like 'Abdi' and 'Asha'. 'Families with Needs' were profiled as receiving "a range of benefits" with names like 'Stacey', while 'Low Income Workers' were typified as having "few qualifications" and were "heavy TV viewers" with names like 'Terrence' and 'Denise'.[20]

[15] https://bigbrotherwatch.org.uk/2018/04/a-closer-look-at-experian-big-data-and-artificial-intelligence-in-durham-police/

[16] Paul Cresswell et al., 'Under the bonnet: Mosaic data, methodology and build', Experian Marketing Services, 1 April 2014. This has since been removed from the Experian website, but we can provide a copy on request.

[17] Mosaic Infographic, Experian, (http://www.experian.co.uk/marketing-services/knowledge/infographics/infographic-new-mosaic.html) Also see Paul Cresswell et al, 'Under the bonnet: Mosaic data, methodology and build', Experian Marketing Services, 1 April 2014, p.7: (http://www.experian.co.uk/assets/marketing-services/presentations/mosaic-data-methodology-and-build.pdf)

[18] https://bigbrotherwatch.org.uk/2018/04/a-closer-look-at-experian-big-data-and-artificial-intelligence-in-durham-police/

[19] https://bigbrotherwatch.org.uk/2018/04/a-closer-look-at-experian-big-data-and-artificial-intelligence-in-durham-police/

[20] https://bigbrotherwatch.org.uk/2018/04/a-closer-look-at-experian-big-data-and-artificial-intelligence-in-durham-police/

5. Durham Constabulary paid £45,913 to Experian for the licensing of their services, including £25,913 for this information,[21] using 'CustodyMosaicCodeTop28', which is described as "the 28 most common socio-geo-demographic characteristics for County Durham",[22] as a predictor in its HART forecasting model, which influenced criminal justice outcomes.

iii. Individual-oriented crime prediction: National Data Analytics Solution (NDAS)

1. The National Data Analytics Solution (NDAS),[23] created by West Midlands Police in partnership with 8 other police forces, including Greater Manchester Police and the Metropolitan Police,[24] is intended to predict serious violent crime using artificial intelligence. The purpose of such predictions is to promote interventions before crimes have been committed. The NDAS was intended for all police forces to use from March 2019, although its operational implementation has been temporarily delayed.[25]

2 . West Midlands Police aims for the system to expand to 34 different use cases (e.g. predicting the likelihood of someone to commit violent crime) for all 44 law enforcement agencies (43 forces including the National Crime Agency). The final product will be "a permanent, cloud-hosted analytics platform" running predictive analytics.[26] West Midlands Police has been given £4,465,000 for the National Data Analytics

[21]Durham PCC Register of Contracts (https://www.durham-pcc.gov.uk/document-library/finance/register-of-contractspcc.pdf)

[22] Sheena Urwin, 'Algorithmic Forecasting of Offender Dangerousness for Police Custody Officers: An Assessment of Accuracy for the Durham Constabulary Model', unpublished thesis, University of Cambridge, 2016, (http://www.crim.cam.ac.uk/alumni/theses/Sheena%20Urwin%20Thesis%2012-12-2016.pdf)

[23] Formerly known as the National Analytics Solution

[24] Founding partners include Greater Manchester Police, Merseyside Police, Metropolitan Police, Staffordshire Police, Warwickshire Police, West Mercia Police, West Yorkshire Police and an unknown (redacted) other.
See: Police Transformation Fund – National Analytics Solution, Final Business Case v6.0 (http://foi.west-midlands.police.uk/wp-content/uploads/2019/01/report1_.pdf)

[25]The West Midlands Police and Crime Commissioner Ethics Committee unanimously voted in favour of being given further information on NAS before it could advise on whether it should go ahead or not.
See: West Midlands Police and Crime Commissioner Ethics Committee, Minutes, 3 April 2019 (https://www.westmidlands-pcc.gov.uk/media/514528/Ethics-Committee-03042019-MINUTES-.pdf)

[26] Founding partners include Greater Manchester Police, Merseyside Police, Metropolitan Police, Staffordshire Police, Warwickshire Police, West Mercia Police, West Yorkshire Police and an unknown (redacted) other.
See: Police Transformation Fund – National Analytics Solution, Final Business Case v6.0 (http://foi.west-midlands.police.uk/wp-content/uploads/2019/01/report1_.pdf)

Solution by the UK Home Office's 'Police Transformation Fund' for 2018/19.[27]

iv. Individual-oriented crime prediction: Offender Assessment System (OASys) and the Offender Group Reconviction Scale (OGRS)

1. The Offender Assessment System (OASys) is a "risk and needs" automated assessment tool, developed jointly by the Prison and Probation Services.[28] It aims to assess the risk of harm offenders pose to others and how likely an offender is to reoffend, as well as assessing offender needs. These risk assessments are used to "target interventions" and to influence the sentence plans given to offenders.[29] An electronic version of the tool was rolled-out across both the prison and probation services, with a new single system being implemented in 2013 through the OASys-R project. By the end of March 2014, almost seven million prison and probation assessments had been collated within the central O-DEAT (OASys Data, Evaluation and Analysis Team) database for over one million offenders.[30]
2. The system collates information on the offenders' previous offences; their education, training and employment; their alcohol and drug misuse; as well as their "attitudes", "thinking and behaviour", "relationships", and "lifestyle". This is done by an assessor who assigns the offender a score based on each category.[31] This data is used alongside the individual's offending record and "offender demographic information" to inform two predictive

[27] Police Transformation Fund – investments in 2018-19 (https://www.gov.uk/government/publications/police-transformation-fund-investments-in-2018-to-2019)

[28] Prison Service Order, Offender Assessment and Sentence Management – OASys (2005) (https://www.justice.gov.uk/downloads/offenders/psipso/pso/PSO_2205_offender_assessment_and_sentence_management.doc); National Offender Management Service, 'A compendium of research and analysis on the Offender Assessment System (OASys) 2009–2013', (2014) (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/449357/research-analysis-offender-assessment-system.pdf)

[29] Prison Service Order, Offender Assessment and Sentence Management – OASys (2005) (https://www.justice.gov.uk/downloads/offenders/psipso/pso/PSO_2205_offender_assessment_and_sentence_management.doc)

[30] National Offender Management Service, 'A compendium of research and analysis on the Offender Assessment System (OASys) 2009–2013', (2014) (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/449357/research-analysis-offender-assessment-system.pdf)

[31] Non-scored categories: Health and other, emotional wellbeing, financial management

algorithms: the OASys General reoffending Predictor v.1 (OGP1) and the OASys Violence Predictor v.1 (OVP1).[32] The Offender Group Reconviction Scale (OGRS) is another static actuarial risk assessment tool used to assess and predict an offender's likelihood of reoffending.[33] The OGRS algorithm uses data on the individual's official criminal history, as well as their age and gender, to produce a risk score between 0 and 1 of how likely an offender is to reoffend within one or two years. There have been several iterations of the OGRS since it was first used in 1996; currently OGRS4 is in use.

3. A 2014 National Offender Management Service analysis found that the OGP1 and OVP1 predictive algorithms generated different predictions based on race and gender. They found that relative predictive validity "was greater for female than male offenders, for white offenders than offenders of Asian, black and mixed ethnicity, and for older than younger offenders".[34] The most sustained differences were by ethnicity, with both OGP1 and OVP1 "working less well for black offenders and OGP1 also working less well for offenders of mixed ethnicity".[35] No assessment of different predictions by ethnicity was carried out in relation to the OGRS4 algorithm. The National Offender Management study from 2014 says there is "a clear need for further studies" to assess, among other things, "whether there are differences… according to age, gender and ethnicity".[36] The recorded disparity in

[32] National Offender Management Service, 'A compendium of research and analysis on the Offender Assessment System (OASys) 2009–2013', (2014) (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/449357/research-analysis-offender-assessment-system.pdf)

[33]https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119184256.ch11

[34] National Offender Management Service, 'A compendium of research and analysis on the Offender Assessment System (OASys) 2009–2013', (2014) (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/449357/research-analysis-offender-assessment-system.pdf)

[35] National Offender Management Service, 'A compendium of research and analysis on the Offender Assessment System (OASys) 2009–2013', (2014) (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/449357/research-analysis-offender-assessment-system.pdf)

[36] National Offender Management Service, 'A compendium of research and analysis on the Offender Assessment System (OASys) 2009–2013', (2014) (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/449357/research-analysis-offender-assessment-system.pdf)

prediction rates between different ethnicities is extremely concerning.

c. Digital evidence

   i. Digital evidence increasingly features in criminal investigations. Police are also using more and more advanced technology to access, download, and analyse digital evidence as part of these investigations.[37] However, technological and legal and policy limitations currently mean that digital evidence collection can be extremely intrusive, obstruct justice, and infringe rights.

   ii. This affects not only suspects but victims of crime, and has become a particular issue for victims of sexual offences. When a complainant indicates that there is digital evidence relevant to a sexual offence on a device in their possession such as a mobile phone, computer or tablet, the devices are typically taken from the complainant and the data extracted. On average, a mobile phone can contain the equivalent of 30,000 A4 pages of documents,[38] ranging through texts, emails, photos, videos, and previously deleted data, and a significant amount of extremely personal and sensitive information. Police also request logins and passwords to victims' social media accounts and personal 'cloud' storage services.

   iii. The out-dated technology in use inevitably leads to disproportionate investigations of victims' digital lives and arguably breaches their privacy rights. The data extraction software police currently use forces the download of everything within a data category, for example all messages or all photos, even if only a single message or photo is needed for evidential purposes.[39] [40] In some cases, police take an entire digital copy of all the information on a device.

   iv. New digital extraction powers in the Police, Crime, Sentencing and Courts Bill claim to regulate this policy area – however, the powers in the Bill are seriously flawed. We are

[37] Privacy International, 'Digital Stop and Search', 27 March 2018 (https://privacyinternational.org/report/1699/digital-stop-and-search-how-uk-police-can-secretly-download-everything-your-mobile)

[38] Dame Vera Baird QC, PCC for Northumbria, 'Letter to Justice Committee', 14 February 2018 (http://www.apccs.police.uk/wp-content/uploads/2016/11/Letter-to-Justice-Committee-Chair-regarding-disclosure-in-criminal-cases-140218.pdf)

[39] https://www.documentcloud.org/documents/4348952-MET-Redacted-Self-Service-Equipment-Kiosk-Local.html in Privacy International, 'Digital Stop and Search', March 2018 (https://privacyinternational.org/sites/default/files/2018-03/Digital%20Stop%20and%20Search%20Report.pdf)

[40] https://www.telegraph.co.uk/news/2018/03/31/police-rolling-technology-allows-raid-victims-phones-without/

concerned that victims and survivors will continue to face choices between protecting their right to privacy and their right to justice. We have written more about what changes are needed to the Bill in written briefings.[41]

v. UK police have previously trialled the use of AI to analyse digital evidence.[42] The Metropolitan Police has confirmed[43] that it has been exploring Cellebrite's 'Analytics Enterprise' artificial intelligence tool, which claims to "detect and match objects within images and video such as weapons, money, nudity and more", use "automatic facial detection", and "analyse links... to reveal hidden connections... and communication patterns".[44]

vi. As this is proprietary technology created by a private for-profit company, there is very little information in the public domain about exactly how the system works or its true capabilities, such as how the system draws such 'links' within communication patterns.

vii. We are extremely concerned that such sensitive police work is being outsourced to experimental systems, with little or no consideration of the myriad transparency, accountability and privacy issues involved. AI analysis is even being trialled to sift through victims' and witnesses' digital information, which is collected in disproportionate volumes. This raises the prospect of a victim of a sexual offence having their digital device and deeply personal information examined and analysed by an experimental, faceless AI system.

viii. Police should not be using artificial intelligence systems to conduct such sensitive investigations.

[41] https://bigbrotherwatch.org.uk/wp-content/uploads/2021/05/Committee-Stage-Briefing-on-digital-extraction-powers-PCSC-Bill-10-NGOs1962.pdf

[42] https://www.theguardian.com/uk-news/2018/may/27/police-trial-ai-software-to-help-process-mobile-phone-evidence

[43] https://www.theguardian.com/uk-news/2018/may/27/police-trial-ai-software-to-help-process-mobile-phone-evidence

[44] https://www.cellebrite.com/en/products/analytics-enterprise/

2. What should new technologies used for the application of the law aim to achieve? In what instances is it acceptable for them to be used? Do these technologies work for their intended purposes, and are these purposes sufficiently understood?
    a. We have answered part of this question throughout our responses in the remainder of the briefing.

3. Do new technologies used in the application of the law produce reliable outputs, and consistently so? How far do those who interact with these technologies (such as police officers, members of the judiciary, lawyers, and members of the public) understand how they work and how they should be used?

   a. **Live Facial Recognition**

      i. The police's facial recognition technology has been incredibly inaccurate, and there are serious problems in general with biased identification rates in facial recognition technologies.

      ii. Our investigations and subsequent report, Face Off: the lawless growth of facial recognition in UK policing, found that the technology was dangerously inaccurate, with facial recognition cameras misidentifying innocent people up to 98% of the time, with an average of 95% of people misidentified.[45]

      iii. A number of independent studies have found that various facial recognition algorithms have demographic accuracy biases – that is, that they misidentify some demographic groups, particularly women and people of colour, at higher rates than white men. A study found that commercial facial recognition technologies, including those created and sold by Microsoft and IBM, had error rates of up to 35% when identifying the gender of dark-skinned women compared to 1% for light-skinned men.[46] A follow up study found that Amazon's 'Rekognition' software mistook women for men 19% of the time, and darker-skinned women 31% of the time.[47]

      iv. The Biometrics and Forensics Ethics Group warned that UK police's use of live facial recognition technology has the "potential for biased outputs and biased decision-making on the part of system operators".[48]

      v. The Metropolitan Police has been aware of these concerns since 2014, when it was raised during an Association of Chief Police Officers (ACPO) 'Facial Imaging Working Group'.[49] We

[45] Big Brother Watch (2018), 'Face Off: the lawless growth of facial recognition in UK policing', 15thMay 2018 (https://bigbrotherwatch.org.uk/wp-content/uploads/2018/05/Face-Off-final-digital-1.pdf)

[46] http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

[47] http://www.aies-conference.com/wp-content/uploads/2019/01/AIES-19_paper_223.pdf

[48] Biometrics and Forensics Ethics Group, Interim report, February 2019 (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/781745/Facial_Recognition_Briefing_BFEG_February_2019.pdf)

[49] Obtained through Freedom of Information Requests.

had asked the police on several occasions between 2017-2019 whether they would carry out or commission demographic accuracy bias testing, and they told us that they would not because they did not view it as an issue.

vi. However, in the Metropolitan Police's written evidence to the Science and Technology Committee in 2019, the force admitted there were issues: "The MPS is cognisant of the concern over the system response with respect to different demographics. We are working to further mitigate potential impact of this within the operational context, where it should be noted, additional checks and balances are in place and the final decision is by a human operator."[50]

vii. In a presentation at University College London on 29 May 2019 about live facial recognition, the Metropolitan Police Senior Technologist, Johanna Morley, admitted that they had found significant gender bias in their technology – that it misidentified women at higher rates than men.[51]

viii. Big Brother Watch has witnessed several incidents that evidence the serious and harmful potential of police live facial recognition misidentifications. At a deployment at Notting Hill Carnival in 2017, we witnessed several innocent women being misidentified as wanted men on the police watchlist. At a deployment in Romford in February 2019, a 14 year old black school child, wearing school uniform, was wrongly identified by the facial recognition system and subsequently surrounded by four plainclothes police officers. He was pulled onto a side street, his arms held, questioned, his phone taken, and fingerprints checked. He was released after ten minutes when police realised the facial recognition 'match' was in fact a misidentification.

b. **Predictive Policing**

i. Geographic Crime Prediction System

1. It has been reported that PredPol has a contractual requirement on customers, including police forces, to engage in promotional activities, such as publicly endorsing PredPol as successfully reducing crime,[52]

[50] Written evidence submitted by Metropolitan Police Service (WBC0005), 19 March 2019: http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/science-and-technology-committee/the-work-of-the-biometrics-commissioner-and-the-forensic-science-regulator/written/97851.pdf

[51] https://www.ucl.ac.uk/jill-dando-institute/events/2019/may/just-looking-learning-police-trials-live-facial-recognition

[52] https://archives.sfweekly.com/sanfrancisco/all-tomorrows-crimes-the-future-of-policing-looks-a-lot-like-good-branding/

despite the lack of clear evidence to corroborate this.[53] The widely claimed benefits of geographic crime prediction are not independently supported by empirical evidence.[54]

2. There are a number of serious issues inherent in the reliability of geographic crime prediction. Such tools use past crime data from police records to predict future crime patterns – but police records represent the crimes, locations and groups that are policed, rather than the actual occurrence of crime. Police data represents systematic under-reporting and systematic over-reporting of certain types of crime and in certain locations,.[55] Police data may represent discriminatory policing practices and societal inequalities, such as those which result in black men being more than 3 times more likely to be arrested than white men in the UK.[56]
3. This means that the data upon which such models are built are not accurate reflections of the true occurrence of crime and are likely to be skewed towards certain crimes and locations, which may reflect social inequalities or discriminatory policing patterns. The 'hotspot' predictions that PredPol creates are also highly targeted, meaning that even small differences in input probabilities lead to huge differences in these output predictions. Resulting predictions are likely to present inaccurate or biased depictions of criminal activity,[57] leading to discriminatory policing interventions.[58]

[53] https://www.techdirt.com/articles/20131031/13033125091/predictive-policing-company-uses-bad-stats-contractually-obligated-shills-to-tout-unproven-successes.shtml

[54] Albert Meijer & Martijn Wessels (2019) Predictive Policing: Review of Benefits and Drawbacks, International Journal of Public Administration (https://www.tandfonline.com/doi/full/10.1080/01900692.2019.1575664)

[55] Lum, Kristian, and William Isaac. 2016. 'To Predict and Serve?' Significance 13 (5): 14–19 (https://rss.onlinelibrary.wiley.com/doi/full/10.1111/j.1740-9713.2016.00960.x); Bennett Moses, L., & Chan, J. (2016). 'Algorithmic prediction in policing: Assumptions, evaluation, and accountability'. *Policing and Society*. (https://www.tandfonline.com/doi/10.1080/10439463.2016.1253695); Barocas, S. and Selbst, A.D., 2016. Big Data's disparate impact. *California law review*, 104, 671. (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899)

[56] Ministry of Justice, 'Black, Asian and Minority Ethnic disproportionality in the Criminal Justice System in England and Wales', 2016 (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/639261/bame-disproportionality-in-the-cjs.pdf)

[57] Innes, M., Fielding, N., & Cope, N. (2005). 'The appliance of science?': The theory and practice of crime intelligence analysis. *The British Journal of Criminology*, 45, 39–57

[58] Albert Meijer & Martijn Wessels (2019) Predictive Policing: Review of Benefits and Drawbacks, International Journal of Public Administration (https://www.tandfonline.com/doi/full/10.1080/01900692.2019.1575664)

4. As such, these predictive models can expand and entrench the biases represented in the crime data,[59] as a result of self-perpetuating 'feedback loops'. This occurs when previous crime data leads to further location-biased predictions, the dispatch of police resources and further crime recording, which is then fed back into the system. Such data-based predictions risk predicting crime and allocating resources in the same areas, creating self-affirming predictions.[60]
5. For example, when policing is disproportionately focused on neighbourhoods with a high black and minority ethnic population, police records will represent higher crime records in those neighbourhoods.[61] There have long been issues with the over-policing of ethnic minorities in the UK, leading to widespread social unrest (for example in St Paul's, Bristol in 1980; Toxteth, Liverpool in 1981; and Broadwater Farm, Tottenham in 1985 and again in 2011.)[62]
6. Multiple studies have found that such geographic crime prediction systems, built and trained using historic police crime records, have lead to self-perpetuating feedback loops particularly in areas with low income and black and ethnic minority populations already subject to excessive policing. This risks reinforcing patterns of inequality.[63] One such study on drug crime in Oakland, California, stated that "locations that are flagged for targeted policing are those that were… already over-represented in the historical police

[59] Lum, Kristian, and William Isaac. 2016. 'To Predict and Serve?' Significance 13 (5): 14–19 (https://rss.onlinelibrary.wiley.com/doi/full/10.1111/j.1740-9713.2016.00960.x)

[60] Ensign et al, (2017) 'Runaway Feedback Loops in Predictive Policing', Cornell University Library, 29 June 2019 https://arxiv.org/abs/1706.0984); Mohler et al (2011), 'Self-exciting point process modeling of crime', Journal of the American Statistical Association (http://www.stat.ucla.edu/~frederic/papers/crime1.pdf)

[61] Custers, B., 2013. Data dilemmas in the information society: introduction and overview. *In*: B. Custers, T. Calders, B. Schermer and T. Zarsky, eds. *Discrimination and privacy in the information society: data mining and profiling in large databases.*: Springer, 3–26. (https://link.springer.com/chapter/10.1007%2F978-3-642-30487-3_1)

[62] Lewis et al, 'Reading the Riots' (2011), London School of Economics and The Guardian, (http://eprints.lse.ac.uk/46297/1/Reading%20the%20riots(published).pdf); Centre for Crime and Justice Studies, 'Policing the riots: from Bristol and Brixton to Tottenham, via Toxteth, Handsworth, etc', (https://www.crimeandjustice.org.uk/publications/cjm/article/policing-riots-bristol-and-brixton-tottenham-toxteth-handsworth-etc)

[63] Ensign et al, (2017) 'Runaway Feedback Loops in Predictive Policing', Cornell University Library, 29 June 2019 (https://arxiv.org/abs/1706.0984); Lum, Kristian, and William Isaac. 2016. 'To Predict and Serve?' Significance 13 (5): 14–19 (https://rss.onlinelibrary.wiley.com/doi/full/10.1111/j.1740-9713.2016.00960.x); Lyria Bennett Moses & Janet Chan (2018) Algorithmic prediction in policing: assumptions, evaluation, and accountability, Policing and Society, 28:7, 806-822 (https://www.tandfonline.com/doi/10.1080/10439463.2016.1253695);

data", and concluded that "allowing a predictive policing algorithm to allocate police resources would result in the disproportionate policing of low-income communities and communities of colour".[64]

7. In addition, such predictive algorithmic models typically lack transparency and accountability. Police officers may not be able to fully understand and interpret the outcomes of predictive models, meaning that predictions can dictate decisions rather than meaningfully inform them. This leads to an accountability deficit, where it is not clear if there is any meaningful decision-making input from police who merely act on predictive algorithms without critical analysis.[65] The use of these systems has the potential to create an unchallengeable narrative of criminal communities.

c. **Individual-oriented crime prediction: Durham Constabulary's Harm Assessment Risk Tool (HART)**

   i. The use of Mosaic geodemographics amounts to discriminatory profiling and is likely to result in inaccurate decisions. The use by police to predict people's "risk", has the potential of affecting potentially life-changing criminal justice decisions. Allowing this kind of profiling data – which includes not only ethnicity data but a whole host of other race and socioeconomic proxy information, including postcodes – to be used in public sector algorithms is discriminatory and, in the criminal justice system, will lead to unjust and inaccurate decisions. This AI risk assessment reinforces existing policing biases and social inequalities, instituting a 'postcode lottery' of justice under the banner of innovation.

   ii. One of the academics instrumental to the development of HART stated to Big Brother Watch verbally that in their opinion the Experian Mosaic data was one of the strongest predictor variables and as such had a valid place in the tool. There is no public data in the available literature to evidence this claim – but even if there were, this statement shows a concerning failure to differentiate between correlation and

[64] Lum, Kristian, and William Isaac. 2016. 'To Predict and Serve?' Significance 13 (5): 14–19 (https://rss.onlinelibrary.wiley.com/doi/full/10.1111/j.1740-9713.2016.00960.x)

[65] Bennett Moses, L., & Chan, J. (2016). Algorithmic prediction in policing: Assumptions, evaluation, and accountability. *Policing and Society*. (https://www.tandfonline.com/doi/10.1080/10439463.2016.1253695)

causation, and treats people for whom such generalised interpretations are not valid as simply collateral. This statistical stereotyping leads to unjust and prejudicial treatment that is the very definition of discrimination.

iii. The HART developers' assessment of the system did indeed recognise that "Some of the predictors used in the model... (such as postcode) could be viewed as indirectly related to measures of community deprivation".[66] They also identified the serious potential for the postcode variables to create 'feedback loops' and reinforce biased criminal justice decisions: "one could argue that this variable risks a kind of feedback loop that may perpetuate or amplify existing patterns of offending. If the police respond to forecasts by targeting their efforts on the highest-risk postcode areas, then more people from these areas will come to police attention and be arrested than those living in lower-risk, untargeted neighbourhoods. These arrests then become outcomes that are used to generate later iterations of the same model, leading to an ever-deepening cycle of increased police attention."[67]

iv. Moreover, Durham Constabulary announced an intention for the HART system to expand beyond the current use alongside Checkpoint, "with the forecasts influencing all of the many other decisions that are made in the wake of bringing a suspected offender into police custody".[68]

v. Following Big Brother Watch's investigation of the HART system, and the use of Experian's Mosaic stereotyping data, we publicised our findings and called for the Experian Mosaic data to be removed immediately (6 April 2018). Durham Constabulary removed the Experian Mosaic data less than three weeks later (24 April 2018).

vi. Separately, since our investigation Experian has also rebranded some of the most crudely titled household profiles in Mosaic, for example changing 'Asian Heritage' to 'Large

[66] Algorithmic risk assessment policing models: Lessons from the Durham Constabulary HART model, M. Oswald, J. Grace, S. Urwin (Durham Constabulary) & G.C. Barnes, 31 August 2017, (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3029345)

[67] Algorithmic risk assessment policing models: Lessons from the Durham Constabulary HART model, M. Oswald, J. Grace, S. Urwin (Durham Constabulary) & G.C. Barnes, 31 August 2017, (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3029345)

[68] Algorithmic risk assessment policing models: Lessons from the Durham Constabulary HART model, M. Oswald, J. Grace, S. Urwin (Durham Constabulary) & G.C. Barnes, 31 August 2017, (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3029345)

Family Living' and 'Crowded Kaleidoscope" to "City Diversity".[69] However, this is a cosmetic change and there is nothing to suggest that the wide range of intrusive underlying data used to create the profiles, including ethnicity data, has changed[70]. Whilst we welcome the removal of overtly offensive stereotype names, we remain deeply concerned about the existence of this profiling data and the role it plays in various areas of public life.

vii. In the US, a similar system to HART called COMPAS, which was also designed to assess the risk of reoffending, was found to be evidencing "significant racial disparities". The COMPAS algorithm is trained on police records, and similarly to the information fed into HART via Mosaic, it uses information on an individuals' education, employment, benefits and financial information. COMPAS routinely underestimated the likelihood of white suspects reoffending, even when the suspect's race was not explicitly included in the dataset. The opposite was true for black suspects who were generally considered at greater risks of recidivism - the system wrongly labelled them as future criminals at twice the rate of white defendants.[71]

d. **Individual-oriented crime prediction: National Data Analytics Solution (NDAS)**

i. Inaccurate outputs from NDAS arise from biased data sources. The police records used to train and predict as part of NDAS include CRIMES;[72] Intelligence Management System (IMS);[73] ICIS;[74] Corvus;[75] Prisoner Intelligence Notification System (PINS);[76] Police National Computer (PNC);[77] OASIS;[78] Drug Intervention Programme (DiP);[79] Organised Crime Group (OCG);[80] and Stop and Search records.[81] West Midlands Police combines data from these 9 police systems, using statistical

[69] https://www.experian.co.uk/assets/marketing-services/brochures/mosaic-ps-brochure.pdf
[70] Paul Cresswell et al., 'Under the bonnet: Mosaic data, methodology and build', Experian Marketing Services, 1 April 2014. This has since been removed from the Experian website, but we can provide a copy on request.
[71] https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing
[72] Records of crimes committed
[73] Police intelligence reports about events, locations and offenders
[74] Custody information
[75] Intelligence, briefing and tasking system
[76] Prisoner information and notification of release
[77] Information on people, crimes, vehicles and property
[78] Event logging system
[79] Drug intervention programme data
[80] Record and mapping of OCGs in the West Midlands Police area
[81] Police Transformation Fund – National Analytics Solution, Final Business Case v6.0 (http://foi.west-midlands.police.uk/wp-content/uploads/2019/01/report1_.pdf); https://www.westmidlands-pcc.gov.uk/media/191164/wmpcc_005_2013_technology_task_force_options_paper_appendix.pdf; https://www.westmidlands-pcc.gov.uk/media/473339/SPCB-05-Dec-17-Item-9-WMP-Change-Portfolio.pdf; =

modelling to identify the strongest 'predictive' fields that indicate someone's likelihood of involvement in a certain crime.[82]

ii. There are serious ethical, data protection and rights issues with several of these data sources. First, the use of data from stop and search, a policing tool that has been consistently used in a biased and discriminatory way, to influence future criminal justice outcomes, will clearly result in similarly biased outcomes. In April 2019, it was reported that black people were 5 times more likely than white people to be stopped and searched in the West Midlands Police area, while Asian people were 2.8 times more likely.[83] In 2017/18, nationally, black people were more than 9 times more likely to be stopped and searched than white people (based on Home Office stop and search data).[84] In May 2019, following the increased use of section 60 'suspicionless' stop and search powers, it was reported that black people were 40 times more likely than white people to be stopped and searched across the UK.[85]

iii. The uncritical general use of crime records within NDAS also embeds biases in policing. As discussed above, police records are not entirely objective and accurate representations of actual criminality and represent societal and structural inequalities as well as recording failures. For example, in 2019, Her Majesty's Inspectorate of Constabulary found that West Midlands Police failed to record more than 16,600 violent crimes each year – 78% of violent crimes and 89% of sexual offences were not recorded when reported.[86] In 2017, HMIC found that West Midlands Police failed to record 38,800 crimes every year – one out of every 6.[87]

iv. These problems are relevant to all police forces. However, it raises particularly serious questions about whether West Midlands Police – or indeed any police - use of data analytics can be credible or fit for purpose when the data they hold is so inaccurate, let alone the fact that police data cannot be

[82] Data Driven Insight & Data Science Capability for UK Law Enforcement (http://www.excellenceinpolicing.org.uk/wp-content/uploads/2017/10/EIP17_2-5_Utilising_Data_Science.pdf)

[83] https://www.westmidlands-pcc.gov.uk/media/514876/SPCB-160419-Item-9a-Stop-and-Search-and-Use-of-Force.pdf

[84] https://www.ethnicity-facts-figures.service.gov.uk/crime-justice-and-the-law/policing/stop-and-search/latest

[85] https://www.theguardian.com/law/2019/may/04/stop-and-search-new-row-racial-bias

[86] https://www.bbc.co.uk/news/uk-england-46867657

[87] https://www.bbc.co.uk/news/uk-england-41178872

considered an accurate record of crime in the first place. In addition, the integration of several 'intelligence' databases (Corvus, IMS) into the NDAS, containing information with potentially questionable or unproven evidential basis, also raises questions about the impartiality and fairness of the system. West Midlands Police have even admitted these problems themselves: "There is potential for bias to be present in the underlying dataset in terms of the recorded incidents of harmful / most harmful offences and within the intelligence reports." [88]

v. West Midlands Police has said that it intends future partners providing data for the NDAS will include the National Health Service, Department for Education, Department for Work and Pensions, Department for Communities and Local Government.[89] The prospect of police or law enforcement basing criminal justice decisions on information from the health service, education, social welfare, local authorities or other public services information is extremely concerning. People should not be profiled based on this information. Such excessive data sharing raises serious ethical and data protection issues and could have a chilling effect on people's access to vital public services.

vi. One of the proposed predictive models evidences further problems inherent in this type of predictive analytics. West Midlands Police developed a predictive risk model, using the police records as above, to identify the 32 strongest 'predictive' fields that indicated someone as an 'influencer' of co-offending.[90] These included the number of times an individual was stopped and searched, the number of intelligence reports about an individual (also analysed above), the number of solo crimes committed by nominal associates, and mentions of the individual in drug habit or addiction records.[91] It is clearly wrong to not only take action against people based on predictions using historic data, but to profile

[88] Police Transformation Fund – National Analytics Solution, Final Business Case v6.0 (http://foi.west-midlands.police.uk/wp-content/uploads/2019/01/report1_.pdf)

[89] Police Transformation Fund – National Analytics Solution, Final Business Case v6.0 (http://foi.west-midlands.police.uk/wp-content/uploads/2019/01/report1_.pdf)

[90] Data Driven Insight & Data Science Capability for UK Law Enforcement (http://www.excellenceinpolicing.org.uk/wp-content/uploads/2017/10/EIP17_2-5_Utilising_Data_Science.pdf)

[91] Data Driven Insight & Data Science Capability for UK Law Enforcement (http://www.excellenceinpolicing.org.uk/wp-content/uploads/2017/10/EIP17_2-5_Utilising_Data_Science.pdf)

and criminalise people based on the actions of others. Recording criminal assumptions about people based on their addictions also raises ethical issues.

vii. There does not appear to be any provision to inform individuals that they have been subject to a NDAS prediction resulting in intervention or whether they will have any opportunity to object to their data being processed or challenge the prediction.

e. **Individual-oriented crime prediction: Offender Assessment System (OASys) and the Offender Group Reconviction Scale (OGRS)**

   i. A 2014 National Offender Management Service analysis found that the OGP1 and OVP1 predictive algorithms generated different predictions based on race and gender. They found that relative predictive validity "was greater for female than male offenders, for white offenders than offenders of Asian, black and mixed ethnicity, and for older than younger offenders".[92] The most sustained differences were by ethnicity, with both OGP1 and OVP1 "working less well for black offenders and OGP1 also working less well for offenders of mixed ethnicity".[93] No assessment of different predictions by ethnicity was carried out in relation to the OGRS4 algorithm. The National Offender Management study from 2014 says there is "a clear need for further studies" to assess, among other things, "whether there are differences… according to age, gender and ethnicity".[94] The recorded disparity in prediction rates between different ethnicities is extremely concerning.

[92] National Offender Management Service, 'A compendium of research and analysis on the Offender Assessment System (OASys) 2009–2013', (2014) (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/449357/research-analysis-offender-assessment-system.pdf)

[93] National Offender Management Service, 'A compendium of research and analysis on the Offender Assessment System (OASys) 2009–2013', (2014) (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/449357/research-analysis-offender-assessment-system.pdf)

[94] National Offender Management Service, 'A compendium of research and analysis on the Offender Assessment System (OASys) 2009–2013', (2014) (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/449357/research-analysis-offender-assessment-system.pdf)

4. <u>How do technologies impact upon the rule of law and trust in the rule of law and its application? Your answer could refer, for example, to issues of equality. How could any negative impacts be mitigated?</u>

    a. Technologies and the Rule of Law

        i. We are concerned that the use of new technologies in the application of the law are engaging human rights and the rule of law in ways that are difficult to assess. The transparency, accessibility and contestability of decision-making processes appears to be largely obstructed by the adoption of commercial technologies, often from private companies, and in any event, are not sufficiently required by law.

        ii. Privacy rights are undoubtedly affected by the trend for ever-more digital governance. Authorities are driven to collect and analyse ever-growing volumes of data about citizens and process it in new and complex ways. However, this privacy shift is not the only way in which rights are affected by the emergence of digital technologies.

        iii. The UK's criminal justice system should embody principles that are at the heart of human rights and rule of law frameworks: equality, transparency, foreseeability, equal and consistent application of the rules, and procedural fairness. These technologies touch on a spectrum of rights: the right to life, the right to health, the right to be free from inhuman or degrading treatment, freedom from discrimination, the rights of children, access to justice, and the right to peaceful enjoyment of property.

        iv. Therefore, decision-making in this context should be transparent, comprehendible to officials and claimants, and challengeable – not just for highly trained lawyers, but for everyone, including disadvantaged and vulnerable people, and people with low levels of digital literacy. This is frustrated both by the nature of the complex technologies in use, and the fact that they are sourced privately and subject to commercial protection. Even the staff using the tools may not know exactly how they work.

        v. Moreover, we believe legal decisions should be human decisions to ensure exercise of official authority is limited, fair and foreseeable in line with the Rule of Law. However, it appears that some decisions are being effectively deferred to automated systems and given merely administrative sign-off by staff. This is, in part, due to ineffective laws.[95]

b. Individual-oriented crime prediction: Durham Constabulary's Harm Assessment Risk Tool (HART)

    i. If the HART system does produce a discriminatory, inaccurate prediction, it is likely to negatively impact the individual and undermine the Rule of Law presumption of innocence and fairness because the system is designed to over-estimate individuals' risk of re-offending: "The HART model intentionally favours... cautious errors, where the offenders' levels of risk are over-estimated".[96]

    ii. This means that the system will predict a "sizeable proportion" of people as being higher risk than they actually are, with the result that innocent people may be incorrectly profiled and subjected to a prosecution they might otherwise have avoided. It's unacceptable that this model deliberately overestimates 'risk' - in effect, the likelihood of guilt – in a way that is fundamentally incompatible with the rule of law and the right to a fair trial. It is vital that individuals are presumed innocent until proven guilty in our justice system.

        1. Durham Constabulary's creation and use of HART exemplifies many of the issues associated with rapid application of algorithms in the justice system: not only profiling, biased feedback loops and discrimination but also data exploitation, de facto automated decision making, and dubious predictions which have consequences for the presumption of innocence and people's right to a fair trial.

c. Individual-oriented crime prediction: National Data Analytics Solution (NDAS)

    i. The NDAS intends to legitimise and support pre-emptive policing interventions using big data analytics and machine learning to make predictions about people's potential future actions in order for police to take action before crimes have been committed. West Midland's Police states that the NDAS will "create meaningful insight and identify value driving patterns which should ultimately lead to crime prediction and

[95] Discussed in relation to Q7.

[96]Algorithmic risk assessment policing models: Lessons from the Durham Constabulary HART model, M. Oswald, J. Grace, S. Urwin (Durham Constabulary) & G.C. Barnes, 31 August 2017, (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3029345)

prevention", enabling police to "make early interventions" and "prevent criminality... by proactively addressing threats".[97] This again raises serious concerns around the presumption of innocence and people's right to a fair trial.

ii. In addition, there is a significant risk of perpetuating and deepening bias as a result of the data used to train the NDAS and to ultimately make predictions. The NDAS uses data about individuals taken from a number of public and private sources. This includes police records, "data ingested from 9 founding partners' source systems", data from other public bodies including social care services, local authorities, education providers and other emergency services, data from private sector organisations and open source data – including social media data.[98] The private sector data includes the use of Experian's Mosaic,[99] considered above.

iii. An independent review of the National Data Analytics System by the Alan Turing Institute Data Ethics Group (ATI DEG) and Independent Digital Ethics Panel for Policing (IDEPP), based only on a draft police report on the NDAS, concluded that there were "serious ethical issues... concerning surveillance and autonomy, as well as the reversal of the presumption of innocence on the basis of statistical prediction".[100]

iv. The reviewers questioned whether it was "ethical to use data in order to intervene for the public good against individuals before they have offended even though this approach will single out individuals who, like the public generally, may not have committed a criminal offence, or who will perhaps not go on to commit a future offence". They also criticised the "reliability or biases in the 'evidence base'" and noted the consequences for "accuracy as well as the legitimacy of preventive action."[101] They stated that the NDAS "seeks to legitimise proactive and preventative policing", "moving law

[97] Police Transformation Fund – National Analytics Solution, Final Business Case v6.0 (http://foi.west-midlands.police.uk/wp-content/uploads/2019/01/report1_.pdf)

[98] Police Transformation Fund – National Analytics Solution, Final Business Case v6.0 (http://foi.west-midlands.police.uk/wp-content/uploads/2019/01/report1_.pdf); See also: West Midlands Police Police and Crime Commissioner Ethics Stakeholder Engagement Proposal (9 March 2018) (Not publicly available but please request a copy if you would like to see it).

[99] Page 14, Police Transformation Fund – National Analytics Solution, Final Business Case v6.0 (http://foi.west-midlands.police.uk/wp-content/uploads/2019/01/report1_.pdf)

[100] ATI DEG and IDEPP, Ethics Advisory Report for West Midlands Police, July 2017 (https://www.turing.ac.uk/sites/default/files/2018-11/turing_idepp_ethics_advisory_report_to_wmp.pdf)

[101] ATI DEG and IDEPP, Ethics Advisory Report for West Midlands Police, July 2017 (https://www.turing.ac.uk/sites/default/files/2018-11/turing_idepp_ethics_advisory_report_to_wmp.pdf)

enforcement away from its traditional crime related role and into wider and deeper aspects of social and public policy."[102]

v. The NDAS evidences many of the issues with predictive analytics, predictive policing, and using historic records to make future predictions. The NDAS not only carries out unacceptable and biased profiling using crude Mosaic data and inaccurate police records - it stigmatises people based on the crimes of others and their social networks. The system's use of biased data and deeply problematic predictors is likely to result in discriminatory feedback loops, reinforcing bias and entrenching structural inequalities. These predictions and the pre-emptive interventions they trigger will result in unfair and unjust criminal justice decisions, reversing the presumption of innocence and possibly infringing people's right to fair trial.

d. Individual-oriented crime prediction: Offender Assessment System (OASys) and the Offender Group Reconviction Scale (OGRS)

i. There does not appear to be any requirement to notify individuals that they have been subjected to this automated risk assessment, nor any mechanism for individuals to challenge the score or the implications it has for their involvement with the criminal justice system. This undermines due process requirements implicit in the Rule of Law.

[102] ATI DEG and IDEPP, Ethics Advisory Report for West Midlands Police, July 2017 (https://www.turing.ac.uk/sites/default/files/2018-11/turing_idepp_ethics_advisory_report_to_wmp.pdf)

5. With regards to the use of these technologies, what costs could arise? Do the benefits outweigh these costs? Are safeguards needed to ensure that technologies cannot be used to serve purposes incompatible with a democratic society?
    a. We have sought to answer this question in the course of our other answers.

6. What mechanisms should be introduced to monitor the deployment of new technologies? How can their performance be evaluated prior to deployment and while in use? Who should be accountable for the use of new technologies, and what accountability arrangements should be in place? What governance and oversight mechanisms should be in place?
    a. A number of serious issues have been identified in this submission in relation to bias, data protection, automated decision-making, and fundamental human rights.
    b. Data is frequently imbued with the prejudices of prior decision makers, and these prejudices will be coded into the decisions of algorithm built using this data.[103] Discrimination can occur because the data being used represent historical patterns of discrimination – and there is no easy method to adjust historical data to rid it of this bias.[104] Even when identifiably biased data is removed from a dataset or algorithm, this does not necessarily remove bias, as other variables can introduce bias into the system by proxy. For example, postcodes are often a proxy for race and socioeconomic status.
    c. There are so many opportunities for bias in data that it has been argued that it is unreasonable to say it can be removed. If it is decided that a system is to be used, developers should at the very least attempt to identify such issues with source datasets, consider their appropriateness, and build tools into models to identify and, if possible, mitigate that bias.[105]
    d. Algorithms being used with significant effect in the public sector should be transparent, with auditable processes and explainable decisions so that they can be understood and challenged by those affected.

[103] Barocas, S. and Selbst, A.D., 2016. Big Data's disparate impact. *California law review*, 104, 671. (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899)

[104] Barocas, S. and Selbst, A.D., 2016. Big Data's disparate impact. *California law review*, 104, 671. (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899)

[105] Professor Suresh Venkatsubramanian presentation at the Amnesty International Expert Meeting on Predictive Policing, 20 May 2019; See also: Friedler, Scheidegger, Venkatasubramanian, Choudhary, Hamilton, Roth (2019). A comparative study of fairness-enhancing interventions in machine learning. In *ACM Conference on Fairness, Accountability and Transparency (*https://arxiv.org/abs/1802.04422)

7. How far does the existing legal framework around new technologies used in the application of the law support their ethical and effective use, now and in the future? What (if any) new legislation is required? How appropriate are current legal frameworks?
    a. Data Protection Act 2018
        i. The Data Protection Act 2018 contains broad exemptions for law enforcement purposes, and as such fails to sufficiently protect citizens' rights – including the right to be free from purely automated decision-making.
        ii. The GDPR safeguards individuals against significant decisions based solely on automated processing.[106] However, the UK's Data Protection Act 2018 makes exemptions from this important GDPR right. Section 14 of the Data Protection Act 2018 permits purely automated decisions with legal or similar significant effects to be made about a subject, in absence of the subject's consent – so long as the subject is notified that the decision was purely automated after the fact. The subject is then to be afforded just one month to request a new decision if they wish.
        iii. However, we are not aware of individuals being notified of purely automated decisions by police, or local authorities, despite the amount of automated-decision-making systems in use as described above.
        iv. This is likely because under section 14 of the Data Protection Act 2018, automated decisions that have significant legal or similar effects on a subject are not necessarily classified as "purely automated" if a human has administrative input. For example, if a human merely ticks to accept and thus enact a serious automated decision, the decision would not need to be classified as "purely automated" under law and as such, the minimal safeguards of notification and re-evaluation would not even apply.
        v. Therefore, welfare and justice decisions could be being made that are for all intents and purposes automated decisions, without individuals being notified of this fact or of their right to appeal. We raised concerns about this during the passage of the (then) Data Protection Bill 2018, which were echoed by the Deputy Counsel to the Joint Committee on Human Rights who said, "There may be decisions taken with minimal human

[106] GDPR, Article 22

input that remain de facto determined by an automated process".[107]

vi. The Data Protection Act 2018 in fact throws open the door for authorities to make significant decisions about people based on big data and automated processing – and weak legal definitions mean that the few safeguards there are may not even apply.

vii. Big Brother Watch believes that two important amendments are required to the Data Protection Act 2018. First, decisions that engage individuals' human rights must never be purely automated decisions; second, automated decisions should be more clearly defined as those lacking meaningful human input.

b. Digital Economy Act 2017

i. Another recent law, the Digital Economy Act 2017 (DEA), undermines privacy in the context of the digital revolution.

ii. Part 5, Chapter 1 of the DEA permits mass data sharing between public authorities and private companies for the improvement or targeting of a public service or benefit provided to individuals or households. Whilst ensuring access to state benefits is a worthy aim, it must be pursued in a proportionate manner and in accordance with data protection law. Critically, this Act lacks a framework for transparency around the data sharing agreements that are made.

iii. Government suggested that the DEA would allow authorities to use bulk data to "identify" and intervene in the lives of "troubled families".[108] This arguably amounts to profiling and risks not only breaching Chapter 3 GDPR, but perpetuating discrimination.

iv. Section 41 DEA further extends the applications of data sharing within and between the state and private companies. Other than fulfilling the purposes for which the data was ostensibly shared, information can be used to prevent or detect crime or anti-social behaviour, for criminal investigations, for legal proceedings, for "safeguarding vulnerable adults and children", for HMRC purposes, or as

[107] Note from Deputy Counsel, 'The Human Rights Implications of the Data Protection Bill', 6 December 2017 (https://www.parliament.uk/documents/joint-committees/human-rights/correspondence/2017-19/Note_Deputy_Counsel_DPBill.pdf)

[108] Digital Economy Bill Factsheet: Better Public Services, Department of Culture, Media and Sport

required by EU obligations. This is a very enabling law that could further institutionalise big data in modern governance and administration. However, the systemic lack of transparency of such data sharing agreements means we know little about how this is working in practice and how people's privacy is being affected.

v. Big Brother Watch recommends that a public inventory of public-private information sharing agreements is established to improve transparency and allow for harmful information sharing agreements to be challenged.

c. Facial Recognition and the Ruling in Bridges

i. The civil liberties campaigner Edward Bridges challenged the trial use of live facial recognition software by the South Wales Police. The force used the technology on 500,000 people on over 60 different occasions. In August 2020, the Court of Appeal ruled that this use breached privacy, data protection and equality laws due to "fundamental deficiencies" in the legal framework.[109]

ii. The Court of Appeal found the governing documents granted an impermissibly wide margin of discretion to individual police officers about who and where live facial recognition could be used.[110] At a minimum, powers relating to new technology must be strictly governed with appropriate framework and safeguards to protect fundamental privacy, equality and data protection rights.

iii. The Court of Appeal also found a breach of the Public Service Equality Duty, stating that all bodies that use novel and controversial technologies such as live facial recognition undertake all reasonable measures to ensure a system is free from bias.[111]

iv. Despite this ruling, the use of live and retroactive facial recognition software by police forces has continued without any transparent and comprehensive framework to ensure the protection of privacy, equality and data protection rights.

v. Big Brother Watch calls for the police to immediately stop using live facial recognition surveillance.

vi. We believe parliament should lead development of legislation for new technologies when it is needed. In the case of live

[109] https://www.libertyhumanrights.org.uk/issue/legal-challenge-ed-bridges-v-south-wales-police/
[110] R (on the application of Bridges) v Chief Constable of South Wales Police [2020] EWCA Civ. 1058, [91].
[111] R (on the application of Bridges) v Chief Constable of South Wales Police [2020] EWCA Civ. 1058, [201].

facial recognition, we believe a legislative ban is urgently needed.

vii. The Government must introduce a requirement for mandatory bias testing of any algorithms, automated processes or AI software used by the police and criminal justice system in decision-making processes.

8. How can transparency be ensured when it comes to the use of these technologies, including regarding how they are purchased, how their results are interpreted, and in what ways they are used?

9. Are there relevant examples of good practices and lessons learnt from other fields or jurisdictions which should be considered?

a. We aimed to answer these questions in our responses above.

10. This Committee aims to establish some guiding principles for the use of technologies in the application of the law. What principles would you recommend?

a. Human-centred: All decisions involving automated processing that engage rights protected under the Human Rights Act 1998 remain ultimately human decisions with meaningful human input.

b. Equality: Introduce a requirement for mandatory bias testing of any algorithms, automated processes or AI software used by the police, criminal justice or administrative system in decision-making processes.

c. Transparency: Introduce a public register of all technologies used in the application of the law and inventory on public-private information sharing agreements.

d. Explain-ability: Algorithmic predictive systems must be transparent with auditable processes and explainable decisions so that they can be understood and challenged by those affected.

e. Prohibit the use of live and retroactive facial recognition surveillance.

f. Prohibit the use of predictive policing systems that have the potential to reinforce discriminatory and unfair policing patterns.

g. Prohibit the practice of indiscriminate data collection and the use of artificial intelligence systems to conduct sensitive investigations.