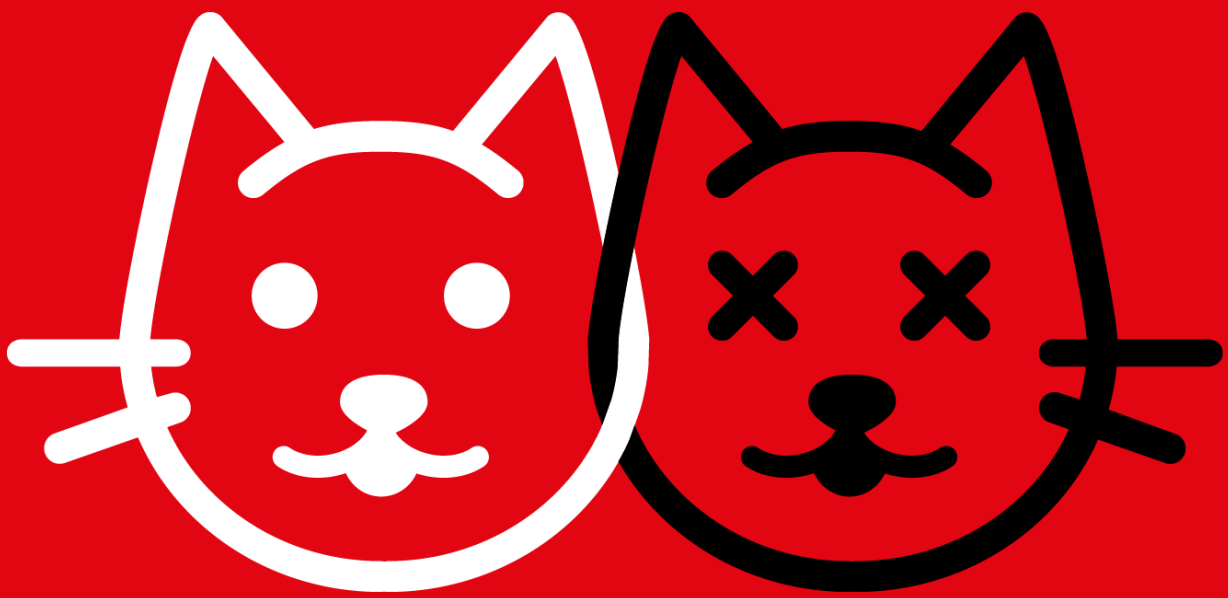
July 2020
APPG AI Evidence Meeting

BIG INNOVATION CENTRE

Face and Emotion Recognition Technologies

How can regulation protect citizens and their privacy?

PARLIAMENTARY BRIEF



Face and emotion recognition technologies: how can regulation protect citizens and their privacy? *is a* Parliamentary Brief based upon the All-Party Parliamentary Group on Artificial Intelligence (APPG AI) Evidence Meeting held online on the 8th June 2020.

This Evidence Meeting was chaired by **Stephen Metcalfe MP** and **Lord Clement-Jones CBE**.

We would like to express our appreciation to the following people for their oral evidence:

- **Professor Nadia Bianchi-Berthouze**, Professor & Deputy Director of UCLIC
- **Dr Temitayo Olugbade**, Research Fellow at UCLIC
- **Matthias Spielkamp**, Co-Founder and Executive Director, Algorithm Watch
- **Dr Seeta Peña Gangadharan**, Assistant Professor in the Department of Media and Communications, LSE
- **Matt Celuszak**, CEO, Element Human
- **Silkie Carlo**, Director, Big Brother Watch
- **Andrew Bud CBE**, CEO and Founder, iProov

Big Innovation Centre is the appointed Secretariat for APPG AI

- CEO, **Professor Birgitte Andersen**
- Rapporteur: **Dr Désirée Remmert**

The video recording of the Evidence Meeting can be found on our websites.

www.biginnovationcentre.com | Email: info@biginnovationcentre.com | @BigInnovCentre
www.appg-ai.org | Email: appg@biginnovationcentre.com | @APPG_AI

PARLIAMENTARY BRIEF

Face and Emotion Recognition: How can regulation protect citizens and their privacy?

All Party Parliamentary Group on
Artificial Intelligence

APPG AI Sponsors

The Group supporters – Blue Prism, British Standards Institution, Capita, CMS Cameron McKenna Nabarro Olswang, Creative England, Deloitte, Dufrain, Megger Group Limited, Microsoft, Omni, Oracle, Osborne Clarke, PwC, Rialto and Visa – enable us to raise the ambition of what we can achieve.

Contents

Introduction

The aim of this APPG AI meeting was to discuss how AI-driven technologies which detect and verify individuals' identity and affective states will change public and private life in the future. Specifically, we wanted to explore how these technologies can be used in a way that protects citizens' privacy and which does not reinforce societal prejudices or exploit vulnerable groups and individuals. We were especially interested in international use cases and research that has been conducted in this field.

The APPG AI Evidence Meeting convened a group of experts in face and emotion recognition research, technologists, journalists, and civil rights activists.:

- **Professor Nadia Bianchi-Berthouze**, Professor & Deputy Director of UCLIC
- **Dr Temitayo Olugbade**, Research Fellow at UCLIC
- **Matthias Spielkamp**, Co-Founder and Executive Director, Algorithm Watch
- **Dr Seeta Peña Gangadharan**, Assistant Professor in the Department of Media and Communications, LSE
- **Matt Celuszak**, CEO, Element Human
- **Silkie Carlo**, Director, Big Brother Watch
- **Andrew Bud CBE**, CEO and Founder, iProov

This meeting was chaired by **Stephen Metcalfe MP** and **Lord Clement-Jones CBE**

APPG AI Evidence Meeting – Global webinar 6th June 2020

Parliament has appointed Big Innovation Centre as the **Secretariat of the APPG AI**, led by **Professor Birgitte Andersen (CEO)**. The Project Manager and Rapporteur for the APPG AI is **Dr Désirée Remmert**.

The expert panel addressed the following questions in their evidence and the subsequent discussion:

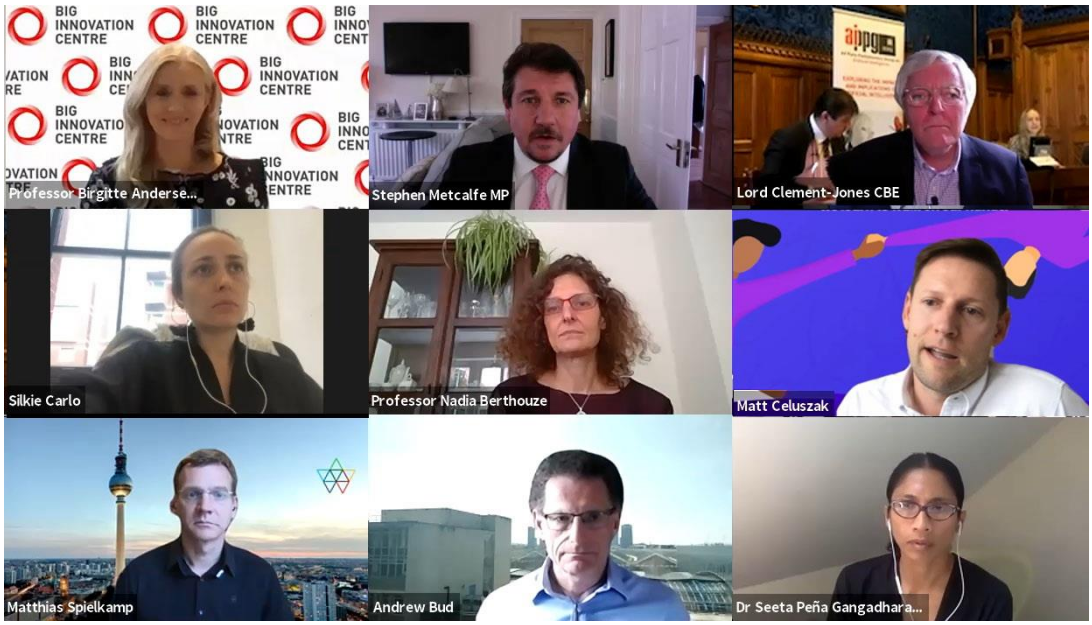
1) **How will face and affect recognition technologies shape law enforcement, the justice system, working environments, and commercial settings in the future?**
2) **Which regulations are needed to guarantee the safe, transparent, and ethical operation of AI-driven face and affect recognition technologies as well as the secure governance of the collected data?**
3) **What can we learn from international use cases of face and affect recognition technologies?**
   – Note, Affect Recognition Technologies (ART) is a concept for technology solutions that claim to detect emotions.

The evidence presented by our expert speakers highlights that **adherence to existing regulation on data privacy, data protection, as well as anti-discrimination laws must be guaranteed** in the deployment of Face Recognition Technologies (FRT)/Affect Recognition Technologies (ART). Further, the evidence points to several critical areas that are not yet sufficiently covered by existing laws and thus require a **clear regulatory framework and oversight body** to secure the safe and fair deployment of Face Recognition Technologies (FRT)/Affect Recognition Technologies (ART) in public and private spaces.

The framework should ensure the **quality and applicability of data sets** used for the training of technologies. It should also **regulate audits and compliance checks**, and outline **rules for the collection, processing, and storage of citizens' biometric data** by Face Recognition Technologies (FRT) and Affect Recognition Technologies (ART) for public and commercial use. **Transparency** about the deployment of these technologies is paramount. The use of Face Recognition Technologies (FRT)/Live Facial Recognition (LFR) in law enforcement and by other public authorities should be accompanied with an **open conversation with the public about the benefits of risks of data-driven surveillance technologies.**

The Parliamentary Brief will first discuss the benefits and challenges in the deployment of face and emotion recognition technologies under consideration of the evidence given by the expert speakers during the APPG AI meeting. It will then explore learnings from international use cases of these technologies and conclude with evidence-based recommendations for policymakers on the deployment of face and emotion recognition technologies and the governance of the collected data. The appendix contains the written evidence of the expert speakers.

1. Benefits and challenges in the deployment of face and affect recognition technologies



APPG AI Evidence Meeting – Global webinar 6th June 2020

The impact of face and affect recognition technologies on various aspects of private and public life has been growing in recent years due to its **expansive use in diverse areas ranging from law enforcement to marketing to recruitment**. It is not surprising that these technologies play an ever more pivotal role in our daily lives as they promise to perform fundamental cognitive tasks - **recognising and interpreting facial features and expressions – better and faster than humans would be able to do**. Facial features and expressions offer us important cues that do not only help us **distinguishing individuals and inferring their emotional states**, but they also **assist us in navigating conversations and non-verbal exchanges**. The importance of the ability to recognise and read faces is fundamental to survival and one of the first capacities that emerge in an infant. Already a few days after birth, infants show a "preference for face-like arrangements that allow the brain to wire itself, with experience, to become expert at perceiving faces" (Feldmann Barrett et al. 2019: 2)[1].

Whereas the **ability to recognise faces and to infer information about affective states from facial expressions is necessary to navigate the social world, it is still unclear to**

[1] Feldmann Barrett, Lisa, Ralph Adolphs, Stacy Marsella, Aleix M. Martinez, and Seth D. Pollak (2019): "Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements." Psychological Science in the Public Interest 20(1) 1-68.

which extent these expressions can be interpreted detached from their specific socio-cultural context. Recent studies of emotional expressions suggest that certain core components of emotions can be recognised cross-culturally. However, it appears to be difficult to reliably infer a more nuanced range of emotions from facial expressions of members outside of one's cultural group (Anger Elfenbein and Ambady 2002: 228; cf Barrett et al. 2019: 46)[2]. Furthermore, the interpretation of emotional expressions is made difficult by a lack of contextualisation. Barrett stresses that

> "the facial configurations in question are not 'fingerprints' or diagnostic displays that reliably and specifically signal particular emotional states regardless of context, person, and culture. **It is not possible to confidently infer happiness from a smile, anger from a scowl, or sadness from a frown**, as much of current technology tries to do when applying what are mistakenly believed to be the scientific facts. Instead, the available evidence from different populations and research domains […] overwhelmingly points to a different conclusion: When facial movements do express emotional states, they are considerably more variable and dependent on context than the common view allow" (ibid. 46).

Emotion recognition technologies, however, mostly lack such necessary contextual information when assessing individuals' affective states by scanning their faces, eye movement, or voice. The issues that arise from the use of non-representative or decontextualised data in the training of emotion recognition technologies[3] are highlighted in the evidence presented by Professor Nadia Bianchi-Berthouze and Dr Temitayo Olugbade from UCLIC:

> "**A unique challenge for Affect Recognition Technologies (ART) is that affect labels are never absolute or objective truth, but rather they are subjective interpretations of either an observer or the subject themselves**. It is critical to ensure that the set of interpretations considered for a given application are ecologically valid and that the predictions of the system are treated as inference, not truth. In obtaining the labels ascribed to the behavioural and physiological data used to develop the system as well as subsequently in developing the system, contextual details that could influence the interpretations of these data should be included. While not trivial to obtain, **it is important to use a balanced and appropriately diverse dataset so as to address individual differences in emotion expression and variations across contexts and groups**."

In order to grasp the nuances of cultural variance in the facial expression and vocalisation of emotions and to prevent assigning labels to stereotypical expressions, so Berthouze and Olugbade argue, training sets must be updated in line with the latest research in this field and

---

[2] Anger Elfenbein, Hillary and Nalini Ambady (2002): "On the Universality and Cultural Specificity of Emotion Recognition: A Meta-Analysis." *Psychological Bulletin* 128(2): 203-235.
[3] See also Sanders, Nada R. and John D. Wood (2020): *The Humachine: Humankind, Machines, and the Future of Enterprise*, pp.153-154. New York and Oxon: Routledge.

must be regularly audited and updated. Further, they highlight, it is important to keep in mind that the results of Affect Recognition Technologies (ART) should be considered inferences and not absolute truths – for this reason, a human in the loop might be necessary to guarantee the accuracy of the findings.

Like Affect Recognition Technologies (ART), the results generated by **Face Recognition Technologies (FRT) are also dependent on the quality of their training data sets**. Whereas Face Recognition Technologies (FRT) are backed with more robust scientific evidence than many affect recognition technologies that are currently employed in HR[4] and commercial settings[5], they come with their **own set of problems that are mostly the result of faulty or non-representative training data, a lack of contextualisation, and a biased selection of attributes**[6]. Recently, especially the use of **Live Facial Recognition (LFR)** technologies (Live Facial Recognition (LFR)) has received critical feedback from the media and human rights activists. Specifically, its deployment for surveillance purposes by police forces has been accused of breaching data protection and anti-discrimination laws.[7]

According to a definition by the Information Commissioner's Office (ICO), **"[l]ive facial recognition (LFR)) technology involves the real time automated processing of digital images containing the faces of individuals, for the purposes of identification, authentication or verification, or categorisation of those individuals"** (ICO 2019: 3). The speakers of the APPG AI evidence meeting agree that transparency in the deployment of Face Recognition Technologies (FRT) and Live Facial Recognition (LFR) as well as informing the public about how Live Facial Recognition (LFR) and Face Recognition Technologies (FRT) work and which data they store are critical to reassure citizens of the lawful, fair, and safe application of Face Recognition Technologies (FRT) and Live Facial Recognition (LFR).

Matthias Spielkamp, founder and executive director of the Berlin based non-profit advocacy and research organisation Algorithm Watch, reports that police forces in numerous European countries already use Live Facial Recognition (LFR) for various purposes. However, Algorithm Watch found that the providers of the deployed technologies did neither guarantee regular auditing of their products, nor were they willing to disclose the training data sets.[8] Data-related issues that impair the accuracy and reliability of the results generated by Face Recognition Technologies (FRT) and Live Facial Recognition (LFR) have been at the centre of the public criticism of these technologies so far, yet Matthias Spielkamp raises a **more fundamental**

---

[4] Nilsson, Patricia (28th February 2018): "How AI helps recruiters track jobseekers' emotions." *Financial Times*.
Accessed 23rd June 2020. https://www.ft.com/content/e2e85644-05be-11e8-9650-9c0ad2d7c5b5
[5] Dupre et al. (2020): "A performance comparison of eight commercially available automatic classifiers for facial affect recognition." *PloS One* 15(4): e0231968.
[6] See Hao, Karen (4th February 2019): "This is how AI bias really happens – and why it is so hard to fix." *MIT Technology Review*. Accessed 23rd June 20202.
https://www.technologyreview.com/2019/02/04/137602/this-is-how-ai-bias-really-happensand-why-its-so-hard-to-fix/
[7] Information Commissioner's Office (31st October 2019): *ICO investigation into how the police use facial recognition technology in public places*. [Report]
[8] See evidence by Matthias Spielkamp in the appendix of this document.

issue concerning the ethical deployment of surveillance technologies which arose in the context of the research:

> "[T]he **accuracy of such systems is not an appropriate criterion to assess whether their use is legitimate or desirable**. […] It is easy to imagine that a 100 percent "accurate" face recognition system could be used to identify members of a certain group, e.g. People of Colour, to subject them to certain measures. **Then the question of the legitimacy of this system resides clearly outside the systems itself** and any kind of technical accuracy (but is of course influenced by claims of accuracy). […] **Is this level of surveillance and scope of centrally stored biometric features the price we are willing to pay as a democratic society in order to, e.g. prevent crime?**"

Likewise, Dr Seeta Peña Gangadharan, assistant professor in the Department of Media and Communications at the London School of Economics and Political Science, alerts to the fact that even highly accurate Face Recognition Technologies (FRT) can do damage if deployed in an inherently flawed system:

> "Much of the technical debate around face recognition has to do with the accuracy of models, including rates of false positives and false negatives. But **the problem of accuracy is an institutional, historical problem, not just a technical one**. Another way stated is: even the most advanced technical improvements can only go so far to make face recognition systems democratic. These systems can only learn so much from their mistakes when introduced to settings where public safety strategies are themselves flawed."

Face Recognition Technologies (FRT) and especially Live Facial Recognition (LFR) in the context of law enforcement and the social justice system have received much criticism from citizens, researchers, and civil rights activists for replicating human biases and perpetuating discrimination. Related technologies like **face verification software**, however, can also be **applied in contexts that improve data privacy and security**. Therefore Andrew Bud CBE, founder and CEO of London-based iProov, advocates for a clear distinction between face recognition and face verification technologies. Other than Live Facial Recognition (LFR), which raises problematic issues around consent and data privacy, face verification technologies appear less controversial from a legal perspective as individual rights as laid out by the GDPR can be more easily observed. **Face verification technologies are using the device webcam to match a face with a certified ID photo. The verification process can be done in realtime, the data does not need to be stored**. Andrew Bud CBE explains:

> "Verification is used to confirm a claim the user makes about who they are. It is not used to identify them. They initiate the process that concludes with the verification, and consent is further assured along the way in the app and user interface. Consent is an absolute legal requirement for normal uses of face verification, according to Article 9 of the GDPR."

However, also **the performance of face verification technologies depends on the quality of the data it has been trained with** and, as Matt Celuszak, CEO of Element Human, points out, they can be **flawed with biases that impair the accuracy of their results**. To retain the safety of face verification technologies, Matt Celuszak asserts, **it is important to make sure that they are used in realtime only**, as a "restricted use case in verification defined by matching capabilities that do not reference a stored database" and that they are **"restricted to device only use cases**".[9]

From the above follows that the accurate and fair deployment of **Affect Recognition Technologies (ART) and Face Recognition Technologies (FRT)/Live Facial Recognition (LFR) is mostly depending on the quality of their training data and the contextualised interpretation of their results**. However, **if deployed in an inherently flawed system, these technologies are likely to compound existing biases and injustices**.

In the following section of this brief, international use cases will be explored in greater detail to obtain insights into the consequences of deploying these technologies on a mass scale.

[9] See evidence of Matt Celuszak in the appendix of this document.

2. What can we learn from national and international use cases of face and affect recognition technologies?

The **deployment of Face Recognition Technologies (FRT) by law enforcement agencies has attracted criticism internationally due to its vulnerability to algorithmic bias and reported cases of misuse and errors based on circumstances and use**. Likewise, vendors and developers of Face Recognition Technologies (FRT) have received negative media attention as they were **accused of aiding coercive political regimes in the monitoring and persecution of citizens**.[10]

One of the most prominent cases that drew **international criticism concerned Microsoft's collaboration with a Chinese military-run university in a research project that explored AI technologies which can be used to create environmental maps by analysing human faces**.[11] China and Russia are currently at the forefront of building complex citizen surveillance systems that rest on the collection and storage of vast amounts of biometric data. This model of **"digital authoritarianism"[12] enables the government to obtain detailed insights into the location, movement, and health of their citizens and thereby monitor their daily lives**. Further, both the Chinese and Russian regimes have been reported to use the current Covid-19 crisis as a cover to expand their surveillance apparatus and to habituate their citizens to AI-driven surveillance technologies in public and private spaces. There is a risk that these technologies will eventually be deployed for other purposes than to control the pandemic.[13]

Whereas most AI technologies that are currently applied in the context of disease control in Western democracies are a far cry from the intrusive population control technologies deployed

---

[10] Human Rights Watch (2019): *China's algorithms of suppression: reverse engineering a Xinjiang police mass surveillance app.* [Report]
Dirks, Emile and James Leibold (2020): *Genomic surveillance: inside China's DNA dragnet.* International Cyber Policy Centre, Australian Strategic Policy Institute. [Report]
Nicola Habersetzer (25th March 2020): "Moscow silently expands surveillance of citizens. *Human Rights Watch.* Accessed on 23rd June 2020. https://www.hrw.org/news/2020/03/25/moscow-silently-expands-surveillance-citizens

[11] Murgia, Madhumita and Yuan Yang (10th April 2019): "Microsoft worked with Chinese military university on artificial intelligence." *Financial Times*. Accessed 23rd June 2020. https://www.ft.com/content/9378e7ee-5ae6-11e9-9dde-7aedca0a081a

[12] Polyakova, Alina and Chris Meserole (2019): "Exporting digital authoritarianism: the Russian and Chinese models." *Foreign Policy at Brookings*. [Report]

[13] (Habersetzer 2020), see footnote no.10 for full reference.
Yang, Yuan et al. (2nd April 2019): "China, coronavirus and surveillance: the messy reality of personal data." Financial Times. Accessed 23rd June 2020. https://www.ft.com/content/760142e6-740e-11ea-95fe-fcd274e920ca
Gebrekidan, Selam (30th March 2020): "For Autocrats, and Others, Coronavirus Is a Chance to Grab Even More Power." *New York Times*.
Accessed 24th June 2020. https://www.nytimes.com/2020/03/30/world/europe/coronavirus-governments-power.html

by certain autocratic regimes, also in the UK concerns, have been voiced about the danger of novel health tech being exploited for sinister purposes in the future[14]. **The growing dependence on data-driven technologies in the current global health crisis, especially those that track people's movements and store biometric data, makes the need for clear regulation of the deployment of technologies that collect citizens' private data even more apparent**.

Recent controversies around the deployment of Face Recognition Technologies (FRT) in law enforcement, however, have mostly been caused by **lack of transparency by vendors and the police about how and where such technologies are applied**. These problems are compounded by **a lack of regulation concerning training data quality and auditing routines of Face Recognition Technologies (FRT), as well as by the inconsistent enforcement of data privacy laws and anti-discrimination laws by the responsible authorities**. Matthias Steinkamp argues that in Poland for instance, automated surveillance used by the police would be hard to detect due to a lack of transparency about the technologies and a low awareness of the issue among citizens. Likewise, in other European states Face Recognition Technologies (FRT) would be deployed in an opaque manner without much regulatory oversight. Matthias Steinkamp reports:

> "In a research conducted in March and April of this year, we [Algorithm Watch] found that at least 11 local police forces in Europe use computer vision to automatically analyse images from surveillance cameras. The risks of discrimination run high, but authorities ignore them."

These findings resonate with the evidence presented by Dr Seeta Peña Gangadharan. She found in her research on the application of Face Recognition Technologies (FRT) in the US and the UK that misclassifications are the result of a complex interplay of social and technical factors and that cultural norms and values play a pivotal role in the collection and preparation of data as well as the creation of algorithmic models. These weaknesses are aggravated by the determinative nature of the data produced by Face Recognition Technologies (FRT). **Determinative data**, Matt Celuszak points out, **is being generated without human verification and is thus entirely dependent on the quality of the technology's algorithmic models**. The application of Face Recognition Technologies (FRT), **thus risks that stereotypes and prejudices ingrained into the model will be exacerbated as the technology is used at scale**. "When problems of false positives and inaccuracy are endemic in society and not only to a technology system" Dr Seeta Peña Gangadharan explains, "the impacts to society as a whole become even more complex."

Nevertheless, a study commissioned by the ICO in 2019 found **strong public support for Live Facial Recognition (LFR) [Live Face Recognition] used for law enforcement**

[14] Vincent, James (5th May 2020): "Without Apple and Google, the UK's contact-tracing app is in trouble." *The Verge*. Accessed 24th June 2020. https://www.theverge.com/2020/5/5/21248288/uk-covid-19-contact-tracing-app-bluetooth-restrictions-apple-google

purposes among British citizens.[15] Yet, the same study also found that citizens were concerned about their privacy and required transparency from authorities about where Live Facial Recognition (LFR) is used and asked for the option of objecting to their images being analysed and stored.

The above national and international examples of the deployment of technologies that use biometric data to analyse a person's location, movement, and physical or mental health demonstrate that is important to weigh the specific dangers these data-driven methods imply against their benefits for the individual and the public.

The evidence presented at the APPG AI evidence meeting illustrates that **the way that Face Recognition Technologies (FRT)/Live Facial Recognition (LFR) are deployed and the results they produce are closely connected to socio-cultural norms and values**. For these reasons, Face Recognition Technologies (FRT) trained with faulty and unrepresentative data risk the exacerbation of pre-existing prejudices and thus aggravate discrimination against individuals and groups. Examples from Russia and China have shown how quickly these technologies can be exploited by coercive regimes in the relative vacuum of international regulation and oversight of these technologies. The recent **withdrawal of Face Recognition Technologies (FRT) products from their use in law enforcement in the US by IBM, Amazon, and Microsoft appears to testify to these tech giants' self-censorship in the absence of national and international regulation for the responsible use of these technologies**.[16]

The above cases have shown that also in Western democracies, a lack of oversight of the training and deployment of Face Recognition Technologies (FRT) risks inaccurate results that can reinforce social injustices. For these reasons, **the speakers at the APPG AI evidence meetings call for reliable measures that ensure the safe and fair deployment of these technologies**.

[15] Information Commissioner's Office (31st October 2019*): ICO investigation into how the police use facial recognition technology in public places*. [Report]
[16] Toulas, Bill (12th June 2020): "Microsoft Joins IBM and Amazon in Facial Recognition Withdrawal." *TechNadu*. Accessed 24th June 2020. https://www.technadu.com/microsoft-joins-ibm-amazon-facial-recognition-withdrawal/104644/
Peter, Jay (8th June 2020): "IBM will no longer offer, develop, or research facial recognition technology." *The Verge*. Accessed 24th June 2020. https://www.theverge.com/2020/6/8/21284683/ibm-no-longer-general-purpose-facial-recognition-analysis-software

3. Policy suggestions for a safe, transparent, and ethical deployment of AI-driven face and affect recognition technologies

The evidence presented by our expert speakers highlights that **adherence to existing regulation on data privacy, data protection, as well as anti-discrimination laws must be guaranteed** in the deployment of Face Recognition Technologies (FRT) and Affect (emotion) Recognition Technologies (ART). Further, the evidence points to several critical areas that are not yet sufficiently covered by existing laws. Therefore, a **clear regulatory framework and oversight body** is needed to secure the safe and fair deployment of Face Recognition Technologies (FRT)/Affect Recognition Technologies (ART) in public and private spaces.

The framework should (i) ensure the **quality and applicability of data sets** used for the training of technologies, (ii) it should **regulate audits and compliance checks**, and (iii) outline **rules for the collection, processing, and storage of citizens' biometric data** by Face Recognition Technologies (FRT) and Affect Recognition Technologies (ART) for public and commercial use. (iv) **Transparency** about the deployment of these technologies is paramount and (v) the use of Face Recognition Technologies (FRT)/Live Facial Recognition (LFR) in law enforcement and other public authorities should be accompanied with an **open conversation with the public about the benefits of risks of data-driven surveillance technologies.**

Towards a Regulatory Framework for Face Recognition Technologies (FRT) and Affect (emotion) Recognition Technologies (ART):

Data quality and contextualisation are essentials plus a "human in the loop"

Data sets used for the training **of Face Recognition Technologies (FRT) and Affect Recognition Technologies (ART) must be representative and relevant for the specific purpose of use** in order to guarantee accurate results that do not replicate inherent biases and thus reinforce societal prejudices and perpetuate systematic discrimination of individuals and groups.

For these reasons, especially the criteria according to which the technologies analyse and structure the collected data (labels/attributes) must be **regularly revised and adjusted to the respective context**. Further, some of these technologies might require a "**human in the loop**", especially those that generate determinative data, to make sure that data is contextualised and free of bias.

Audits and compliance checks must be conducted:

Face Recognition Technologies (FRT) and Affect Recognition Technologies (ART) must be regularly audited to guarantee their safety and to check if they are still fulfilling the desired purpose. An **oversight body should certify the functionality of the technology** as well as the **correct processing and storage of the collected data**. Further, **algorithm auditors should be implemented to alert to algorithmic bias and overfitting**. **Compliance checks should be performed** by public authorities to guarantee that the technologies are deployed lawfully. Fundamental rights impact assessments can also contribute to identifying potential biases.

Transparency must be present:

Public and private institutions and businesses **must be transparent about where Face Recognition Technologies (FRT) and Affect Recognition Technologies (ART) are deployed, what kind of data they collect, and how it is processed and stored**. Further, there should be transparency about the data sets used for the training of the deployed technologies as well as about the **assigned labels** in the analysis of the collected data. Moreover, there should be a **public conversation on the democratic implications of face recognitio**n and methods on how to guarantee their safe and fair deployment in public spaces.

Raise awareness of the distinctive functions and risks of Face Recognition Technologies (FRT)/Affect Recognition Technologies (ART):

The **different types of face and affect recognition technologies come each with specific sets of risks that should be considered in the design of a regulatory framework**. Further, Face Recognition Technologies (FRT) and Affect Recognition Technologies (ART) are **built upon scientific findings** of varying robustness and thus the data they generate might need to be **assessed by a human for contextualisation and accuracy**. Further, **there exist large discrepancies** in the needs for processing and storage of data between these technologies, **especially between face recognition and face verification technologies**, which might affect their applicability depending on the specific purpose and context.

The suggested regulatory framework should also consider the **challenges that arise** within the context in which the technology is deployed, especially whether it is deployed for public or commercial use. The GDPR is currently still regulating data governance in public and commercial use. However, the implementation of **an upgraded regulatory framework for the post-Brexit era is pressing to prevent complications in the national and international application of Face Recognition Technologies (FRT) and Affect Recognition Technologies (ART) by British companies and public institutions**.

5. Evidence

Professor Nadia Bianchi-Berthouze, Professor & Deputy Director of UCLIC and Dr Temitayo Olugbade, Research Fellow at UCLIC

APPG AI Evidence Meeting – Global webinar 6th June 2020

with contribution from

APPG AI Evidence Meeting – Global webinar 6th June 2020

1. **How will affect recognition technologies (Affect Recognition Technologies (ART)) shape law enforcement, the justice system, working environments, and commercial settings in the future?**

Affect is critical to human life. It interacts with our cognitive processes [1] and physical activity [2] and facilitates human-human communication. As such affect recognition technology (Affect Recognition Technologies (ART)) capable of reading facial and vocal expressions, body movement, tactile behaviour or physiological responses can better address people's needs, preferences, and intentions and spersonalise the services and support that they provide. Here a few examples of the several possibilities that exist:

Better working environment: Stress level automatic recognition could help people better sorganise their working patterns. Human resources departments could use anonymous staff data to improve working patterns and environment to support well-being and productivity. e.g. in HUMAN (humanmanufacturing.eu), we investigated how technology could adapt support to factory floor workers by providing just-in-time information based on detected levels of stress, confusion, or frustration with respect to more objective information (e.g. phase of the design, errors, etc.). Detection of pain, anxiety, and fatigue could help spersonalise support provided by robots/exoskeletons/wearables (HUMAN/EMOPAIN: www.emo-pain.ac.uk) as needed so that robot technology collaborates with rather than replace human workers. Affective technology could also help smooth social interactions in working environments where stress, pressure, lack of colocation, or even limited affective skills [3] may hinder empathy and collaboration.

Commercial environment: There is a lot of waste due to lack of understanding of consumers' needs. Technology that detect consumers' affect can help complement limited self-report-based approaches to evaluate peoples' responses to products. A salesperson would observe their customers as they engage with, touch, or try out products. Similarly, technology that reads facial and body expressions or even affective touch [4][5][6] and physiological responses can get a deeper insight into how their customers feel or what their preferences are. Social robots & conversational agents could become possible extension of a salesperson in more private contexts (e.g., changing room) and even help boost self-esteem. Such in-store technology could also enable crowdsourcing of customer experience in the physical shop and share them with online customers for personalisation of feedback, suggestion, and support.

2. **Which regulations are needed to guarantee the safe, transparent, and ethical operation of AI-driven face and affect recognition technologies as well as the secure governance of the collected data?**

Existing regulations (e.g. the Data Protection Act 2018) need to be enforced in the use of Affect Recognition Technologies (ART)s and it is critical to further ensure that these technologies are not misappropriated, i.e. that they are only used within the limitations of their capabilities, and also that they are used lawfully, fairly, and transparently. There are additional standards that need to be met to foster appropriate development and application of these technologies.

Fit for purpose: Affect Recognition Technologies (ART)s must be fit for purpose. Affect Recognition Technologies (ART)s have mainly been investigated from a machine learning perspective and sometimes based on outdated models of emotional expressions [7] or models that do not fully reflect the context of use. This controlled approach was critical to pioneer the field. Now, it is important that the development of Affect Recognition Technologies (ART)s is based on the use application perspective, bringing together AI and HCI/UX experts as well as experts and users from the context domain [8]. Affect Recognition Technologies (ART)s must for example i) address the affective states that matter for the given application rather than just basic emotions (e.g., embarrassment rather than simply sadness); ii) be trained to recognise the variety of everyday expressions and not stereotypical expressions; iii) be based on the most relevant affective cues, which may not be facial expressions; iv) consider the contexts of use as affective expressions are affected by social rules, ongoing activity, etc.

Datasets: The dataset used to then build an Affect Recognition Technologies (ART) is a principal element that defines how well it performs and the limits to its inference. A unique challenge for Affect Recognition Technologies (ART) is that affect labels are never absolute or objective truth, but rather they are subjective interpretations [9] of either an observer or the subject themselves. It is critical to ensure that the set of interpretations considered for a given application are ecologically valid and that the predictions of the system are treated as inference, not truth. In obtaining the labels ascribed to the behavioural and physiological data used to develop the system as well as subsequently in developing the system, contextual details that could influence the interpretations of these data should be included [7][10]. While not trivial to obtain, it is important to use a balanced and appropriately diverse dataset so as to address individual differences in emotion expression and variations across contexts and groups (e.g. [11]). To further ensure the validity of the created system for the given use case scenarios, it is essential that it is rigorously tested using established techniques and on representative real-life examples in those scenarios. Certain applications may demand regular evaluation even after deployment, to account for system updates and detect unexpected system behaviours (e.g. when the system comes across new data that is not represented in the dataset that the system was built with).

3. What can we learn from international use cases of face and affect recognition technologies?

Need for transparency: It is clear from debates (e.g. [12]-[14]) on the use of AI technologies in other countries and what we know from other UK case studies (e.g. [15]) that there needs to be a meaningful discussion with the public. In the first instance, its aim would be to fully capture pertinent concerns and gain an understanding of what systems need to be put in place to manage risks before Affect Recognition Technologies (ART) can be satisfactorily adopted. Such discourse should also include emotion science researchers, affective computing scientists, technology developers, relevant regulation bodies, and the various stakeholders. Beyond this, the public needs to be continually engaged in the conversation on the use of this technology. Transparency is one of the tools necessary for facilitating such involvement [16] and build trust with users. This includes clearly informing subjects about how their data is evaluated and what the evaluation is being used for (as is required by the Data Protection Act

2018), for deployed technology. It is equally valuable for those that provide Affect Recognition Technologies (ART) as well as companies or agencies that apply them to publish publicly accessible information about the datasets used to build their technology and the evaluations conducted to validate it [17].

[1] Bandura, A. (1977). Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215.

[2] Vlaeyen, J.W.S., Morley, S., Crombez, G. (2016). The experimental analysis of the interruptive, interfering, and identity-distorting effects of chronic pain. *Behaviour Research and Therapy*, *86*,23–34.

[3] Voss, C., Schwartz, J., Daniels, J., Kline, A., Haber, N., Washington, P., Tariq, Q., Robinson, T.N., Desai, M., Phillips, J.M., & Feinstein, C. (2019). Effect of wearable digital intervention for improving socialization in children with autism spectrum disorder: A randomized clinical trial. *JAMA Pediatrics*, 173(5), 446–454.

[4] Gao, Y., Bianchi-Berthouze, N., & Meng, H. (2012). What does touch tell us about emotions in touchscreen-based gameplay? *ACM Transactions on Computer-Human Interaction, 19* (4), 1–31.

[5] Atkinson, D., Orzechowski, P., Petreca, B., Bianchi-Berthouze, N., Watkins, P., Baurley, S., . . . Chantler, M. (2013). Tactile Perceptions of Digital Textiles: a design research approach. CHI'13, 1669-1678.

[6] Petreca, B., Baurley, S., & Bianchi-Berthouze, N. (2015). How do designers feel textiles? *International Conference on Affective Computing and Intelligent Interaction*, 982–987

[7] Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, *20*(1), 1–68

[8] Olugbade, T., Singh, A., Berthouze, N., Marquardt, N., Aung, M. S. H., & Williams, A. (2019). How can affect be detected and represented in technological support for physical rehabilitation? A*CM Transactions on Computer-Human Interaction*, *26*(1), 1–29

[9] Blakemore, S. & Decety, J. (2001). From the Perception of Action to the Understanding of Intention. *Nature Reviews Neuroscience*, *2*(August), 561–567

[10] viezer, H., Trope, Y., & Todorov, A. (2012) .Body Cues, Not Facial Expressions, Discriminate Between Intense Positive and Negative Emotions. *Science, 338*(6111):1225–1229

[11] Alghowinem, S., Goecke, R., Epps, J., Wagner, M., & Cohn, J. (2016) Cross-Cultural Depression Recognition from Vocal Biomarkers. *Interspeech,* 1943–1947

[12] ACLU Masschusetts. (2019, 18 June). Massachusetts Voters Strongly Support Pausing Use of Unregulated Face Recognition Technology. ACLU Massachusetts. Retrieved from https://www.aclum.org/en/news/massachusetts-voters-strongly-support-pausing-use-unregulated-face-recognition-technology)

[13] House Committee on Oversight and Reform. (2020, 15 January). Hearings: Facial Recognition Technology (Part III): Ensuring Commercial Transparency & Accuracy. House Committee on Oversight and Reform. Retrieved from https://oversight.house.gov/legislation/hearings/facial-recognition-technology-part-iii-ensuring-commercial-transparency

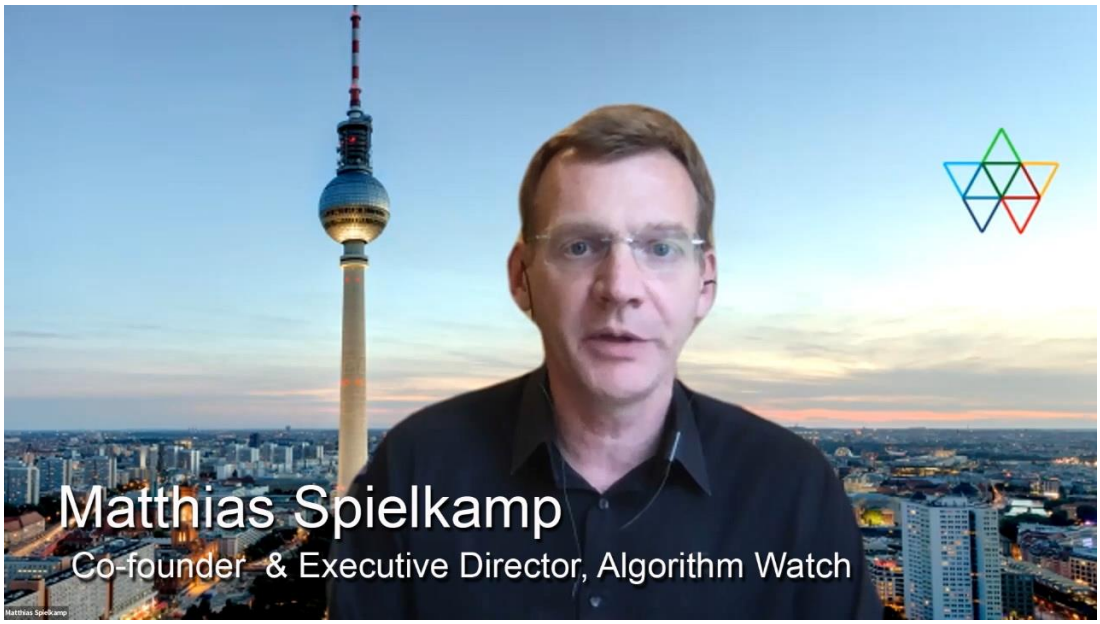
[14] Vox Media (2019, 5 December). How 2020 Democratic candidates would regulate facial recognition. Vox. Retrieved from https://www.vox.com/policy-and-politics/2019/12/3/20965470/2020-presidential-candidates-facial-recognition

[15] Stilgoe, Jack. (2007). The (co-) production of public uncertainty: UK scientific advice on mobile phone health risks. *Public Understanding of Science*, *16*(1), 45–61

[16] Harrison, T.M., Guerrero, S., Burke, G.B., Cook, M., Cresswell, A., Helbig, N., Hrdinova, J., & Pardo, T. (2012). Open government and e-government: Democratic challenges from a public value perspective. *Information Polity*, *17*(2), 83–97

[17] Arnold, M., Bellamy, R.K., Hind, M., Houde, S., Mehta, S., Mojsilović, A., Nair, R., Ramamurthy, K.N., Olteanu, A., Piorkowski, D., & Reimer, D. (2019). FactSheets: Increasing trust in AI services through supplier's declarations of conformity. *IBM Journal of Research and Development*, *63*(4/5), 6:1–6:13

Matthias Spielkamp, Co-Founder and Executive Director, Algorithm Watch



APPG AI Evidence Meeting – Global webinar 6th June 2020

The question of how face recognition systems should be regulated remains open. Options include establishing a moratorium to gain time for further deliberation, banning their use for certain purposes, or even entirely.

The current situation:

It is impossible to provide an overview of the use and implications of face recognition systems globally, so I will focus on results of Algorithm Watch's own research and combine it with recent developments in the field. In research we conducted in March and April of this year, we found that at least 11 local police forces in Europe use computer vision to automatically analyse images from surveillance cameras. The risks of discrimination run high but authorities ignore them. Pedestrians and motorists in Cannes, Marseille, Nice, Nîmes, Roubaix,Toulouse and Yveline (France), Brussels and Kortrijk (Belgium), Prague and Prostějov (Czech Republic), Warsaw (Poland), Mannheim (Germany), and Marbella (Spain) are constantly monitored for abnormal behaviour. Police in these cities connected the video feeds of surveillance cameras to automated systems that claim to detect suspicious movements such as driving on bus lanes, theft, assault or the coalescence of aggressive groups. All automated surveillance techniques in use in the cities listed above rely on Machine Learning. This approach requires that software developers feed large amounts of scenes depicting normality, and others representing the situations considered abnormal, to computer programs. The programs are then tasked with finding patterns that are specific to each type of situation.

Spurious correlations:

Machine Learning has many applications that are now routinely used, such as reverse image search or automated translation. But the drawbacks of this technique are well known. The software does not understand a situation in the human sense, it only finds inferences in the data it has been given. This is why, after decades of controversy, Google Translates still renders the gender-neutral “they are doctors” in German as “sie sind Ärtze” (masculine) and “they are nurses” as “sie sind Krankenschwestern” (feminine). Google Translate was not programmed to be sexist. The corpus of texts it received happened to contain more instances of male doctors and female nurses. What is true of automated translation is true of automated image recognition, known as computer vision. On 7th April, AlgorithmWatch revealed that Google Vision, an image labelling service, classified a thermometer as a “tool” in a hand that had a light skin tone, and “gun” in a dark-skinned one. (Google since changed their system). Results provided by Google Vision Cloud before 6th April.

Misconceptions:

AlgorithmWatch asked several vendors of computer vision solutions to police forces what training data they used, and how they ensured that their programs were not discriminatory. A spokesperson for BriefCam, which is used by police forces from Warsaw to Roubaix, stated

in an email that, because the software did not use skin tone as a variable, it could not discriminate. This is a commonly held misconception. Machine Learning software is designed to find patterns that are not specified by their programmers in order to achieve their results. This is why Google Translate produces sexist outcomes, and Google Vision produces racist outcomes, although they were not explicitly programmed to take into account gender or skin tone. BriefCam’s spokesperson added that they used “training datasets consisting of multi-gender, multi-age and multi-race samples without minority bias,” but declined to provide any evidence or details. The police of Etterbek, in Brussels, uses computer vision to automatically spot illegal trash disposal. A spokesperson for the city wrote that the system did not take skin tone or any other individual trait into account but failed to provide any information about the training data set their software was built on. A spokesperson for Frauenhofer IOSB, which powers the automated surveillance of Mannheim, Germany, claimed that their software could not be discriminatory because it relied on 3-dimensional modelling of body shapes. It analysed movements, not images, and therefore did not use skin tone, he added. Details on the training data set and its diversity was not provided. Avigilon declined to comment. One Télécom, Two-I and Snef did not reply to numerous emails.

Invisible issue:

Automated surveillance is hard to detect. Police forces have no obligation to disclose that they use it and the calls for tenders are rarely published. In Poland, for instance, AlgorithmWatch was told that any information on the issue was “confidential”. The details of their automated surveillance operation were only available in an article in their internal publication, Police Magazine – which is available online. This invisibility makes it hard for civil society

sorganisations to weigh in. AlgorithmWatch spoke to several anti-discrimination sorganisations at the local and national level. While their spokespersons acknowledged the importance of the issue, they said they could not address it for lack of awareness among the population and for lack of monitoring tools. Meanwhile, automated surveillance has the potential to dramatically increase discriminatory policing practices.

Unaudited:

How much automated surveillance impacts discrimination in policing is not known. None of the vendors or cities AlgorithmWatch contacted conducted audits to ensure that the output of their systems was the same for all citizens. Nicole Romain, spokesperson for the Agency for Fundamental Rights of the European Union, wrote that any institution deploying such technologies should conduct a “comprehensive fundamental rights impact assessment to identify potential biases”. When it came to computer vision in policing, she was not aware that any such assessment had ever been made.

What can we learn from these use cases?

Accuracy is not an appropriate measure. What seems paradoxical but is important to understand is that the accuracy of such systems is not an appropriate criterion to assess whether their use is legitimate or desirable. A common response to individuals and organisations scriticising the use of face recognition systems is that their accuracy will improve with the size of training data sets and progress in algorithm development. This has been claimed for decades, and it has so far turned out to be a false claim that in many cases was most surely consciously made strategically in order to soothe criticism. I will not argue in detail why there are very good reasons to believe that there is a very slim chance the claim will ever be true. Because it is beside the point, just imagine if we were capable of developing a system that was 100 percent “accurate” in the sense that it would correctly identify all humans strolling down high street whose images are stored in a database. Then we still have to answer two questions:

1) Is the purpose of the system legitimate? It is easy to imagine that a 100 percent “accurate” face recognition system could be used to identify members of a certain group, e.g. People of Colour, to subject them to certain measures. Then the question of the legitimacy of this system resides clearly outside the systems itself and any kind of technical accuracy (but is of course influenced by claims of accuracy).

2) Is this level of surveillance and scope of centrally stored biometric features the price we are willing to pay as a democratic society in order to, e.g. prevent crime?

Where to go from here?

In January, a preliminary version of the European Commission’s White Paper on AI was leaked to the media. This draft argued for “a time–limited ban on the use of facial recognition technology in public spaces”, detailing that “use of facial recognition technology by private or

public actors in public spaces would be prohibited for a definite period (e.g. 3–5 years) during which a sound methodology for assessing the impacts of this technology and possible risk management measures could be identified and developed." Not surprisingly, this recommendation never made it into the final paper, which only announced that "the Commission will launch a broad European debate on the specific circumstances, if any, which might justify such use, and on common safeguards."

Not only that: In an apparently diametrically opposed move, the European Union's ICT agency for internal security and border control, eu-LISA, last week "signed a framework contract for a new biometric matching system which aims to create a database of fingerprints and facial images of more than 400 million third country nationals by 2022." It may seem odd and even naive, but to me the discussions that happened inside the European Commission in the drafting of the White Paper contains a heartening aspect.

Dr Seeta Peña Gangadharan, Assistant Professor in the Department of Media and Communications, London School of Economics and Political Science

APPG AI Evidence Meeting – Global webinar 6th June 2020

My comments focus on putting the debate on face recognition into the larger context of equality and justice. I have been researching data and discrimination for the last ten years. I speak with reference to my own researchbut also in relation to other social science and computer science research which informs my work. Much of the technical debate around face recognition has to do with accuracy of models, including rates of false positives and false negatives. But the problem of accuracy is an institutional, historical problem, not just a technical one.

Another way stated is: even the most advanced technical improvements can only go so far to make face recognition systems democratic. These systems can only learn so much from their mistakes when introduced to settings where public safety strategies are themselves flawed. Let me unpack this.

Computer scientists point to two types of problems regarding accuracy in machine learning. The first set of problems pertains to flawed data and flawed models. Face recognition systems misclassify data because they have been trained on flawed data or something shifts in the training data that then causes a shift in a model. Face recognition systems can also misclassify data because they have flawed models. For example, MIT researchers demonstrated that face recognition algorithms misclassify female faces of colour at higher rates than other faces.
The second problem pertains to attacks. Face recognition systems might be attacked in ways that interfere with their ability to detect the presence of face, then extract the face, and then

srecognise it. For example, researchers at Carnegie Mellon University found that amending a face—through the addition of certain kinds of eyewear, or changes in lighting or camera angle—interfere with the accuracy rates of face recognition systems.

What I take from this research is a chilling observation: false positives are a problem from data to model. Whether intentionally or unintentionally, a system can be compromised at many different points in the process of machine learning. The problem which leads to inaccuracies might be due to the persons or institutions responsible for collecting certain kinds of data, which later gets used to train models. It might be the persons or institutions that decide to use a particular training dataset. The problem of inaccuracies might stem from a system designer or designers who create a flawed algorithmic model, or the institution that gives it the green light. It might be someone (or something) outside the system who interferes with system performance. Problems of inaccuracy might arise simply due to the course of human history changing and causing shifts in a training dataset.

Now graft this technical system onto a social system, with entrenched cultural norms and values, which has its own problems of accuracy. Misclassification in a sracialised society is also known as racial profiling. If you are a particular race or ethnicity, you are misidentified as criminal, because of a deep-seated notion that members of Black and minority ethnic populations are dangerous and commit more crimes. Being a false positive is a default existence in Black and Brown communities. And the problem of false positives—i.e. racial profiling—will be exacerbated by the widescale adoption of face recognition technology.

Recall some facts about recent analyses of police data and racial profiling:

A joint study by Liberty and the Guardian found that 22% of Black and minority ethnic

- people received fines in the UK, though they represent 15.5% of the total population.
- Another analysis done for the Guardian completed by University College London lecturer

Dr Kristian Pozsch showed that Black people made up 31% of arrests in London during lockdown, despite making up 12% of London's population.

Additionally, recall facts about face recognition pilot programs.

The use of face recognition during Notting Hill Carnival in 2017, which resulted in 35 false

- matches and one erroneous arrest.6
- According to University of Essex researchers, trials in 2019 for face recognition in London relied on watch lists that were not current. The same trials yielded 8 correct identifications out of 42.7

When problems of false positives and inaccuracy are endemic in society, and not only to a technology system, the impacts on society as a whole become even more complex. In the research that I co-lead, we have found that the surveillance works against public safety. It

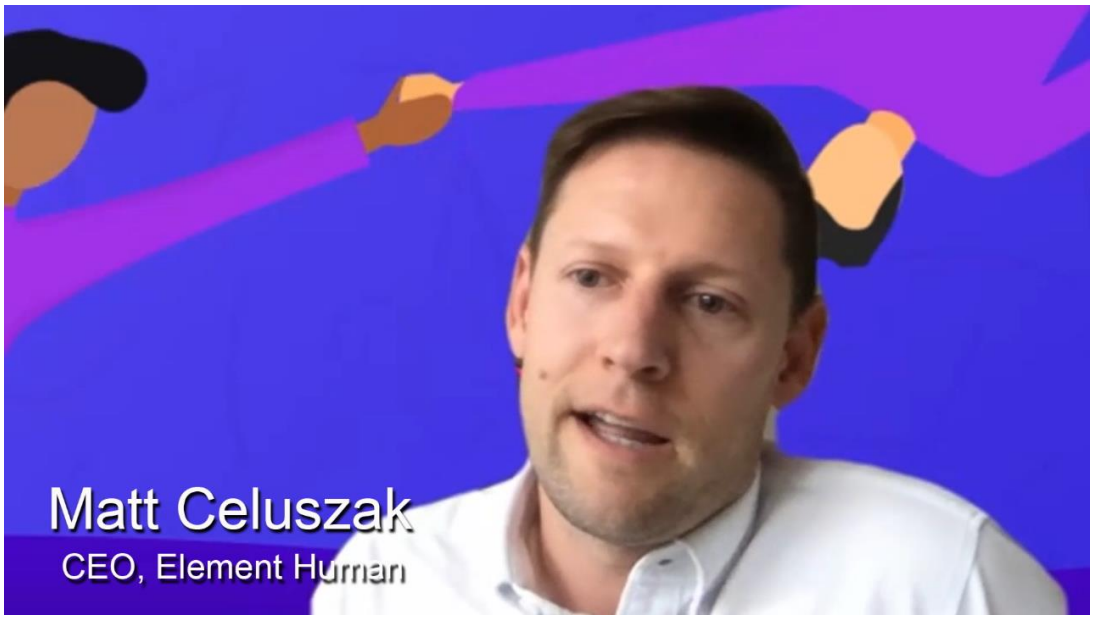
compounds mistrust in those institutions meant to defend and protect public safety. Surveillance contributes to a false sense of security. Any measures that attempt to solve inaccuracies in a technical system are simply doomed to fail in such a context.

Many of the solutions that flow from our research are non-technical ones:

- They concern measures to stop racial profiling.
- They include setting standards for the integrity of data collected about police misconduct and violence.
- They involve the rectification of flawed criminal records.
- They pertain to making it easier to know what democratic safeguards are being put in place in contracts between states and technology vendors.
- They involve having frank and ongoing conversations about the need for such surveillance technologies.

The biggest innovation that the government might support now is one such frank conversation on democratic implications of face recognition. That conversation should happen now, and it should happen in a way that asks people whether the costs of surveillance are the costs that we are willing to bear.

Matt Celuszak, CEO, Element Human



APPG AI Evidence Meeting – Global webinar 6th June 2020

Face technologies are here to stay. They need work, clarity, and regulation.

Disclaimer: I am not a god at machine learning, nor am I a guru psychologist. I'm not going to sit here and tell you that facial recognition is the bee's knees and you should throw away your privacy to protect your freedom. Nope, I'm just the guy who couldn't read body language to save his life. Too many awkward, face-palm moments is a testament to that. I am a glutton for punishment. So, I started CrowdEmotion to see if technology could help me understand people better.

Background: We spent 7 years curating over 2.5 Billion facial attention and emotion data points with consent from 450,000 people in 89 countries to inform 40 businesses to understand how body language relates to human behaviour and business performance. In the process, we built, and rebuilt, emotion recognition technology 5 times to solve for a number of both technical and systematic bias challenges prevalent within the technologies on the market.

There are technical and societal problems that need to be solved, but after 7 years of seeing the benefits of face detection technologies, they are worth solving. Carefully.

Below we explore what the different technologies are, their use cases and the considerations needed for policy and decision making.

Defining Verification, Recognition, and Biometrics

1) Facial verification:

Typically used to identify an individual in realtime at the point of interaction, it uses the device webcam to take a picture of your face and match it to a photo of your face from an accepted ID (like a passport, drivers license, residence permit etc). The photo can often be uploaded or can be simply displayed to the device webcam as well.

Pros:

- Directly identifies an individual
- Protects consumers from fraudulent access to their accounts and transactions
- Can be done in real-time with no need to store identity data long term

Cons:

- Ability to use can be limited by poor quality webcams, low lighting and racial bias
- Data is determinative

2) Facial Recognition:

Like a fingerprint, facial recognition is typically used to identify individuals and link information to their digitally stored profiles often associated with government ID. It works by using cameras to capture images of your face and match it to a database of profiles to start building a deep characteristic of you as an individual. While it can be used to verify the identity of individuals like Facial Verification, it's primary use is within security and law enforcement to identify, track and source criminals through camera networks.

Given the law enforcement use and the technological bias prevalent within computer recognition systems today, this is where the distinct issues arise as citizens are not empowered to dispute the veracity of the technology with their tax-funded law enforcement agencies putting a relatively large power imbalance in the hands of authorities.

Pros:

- Directly Identifies an individual
- Can track criminals through identity
- Fast track check-in processes
- Create personalised experiences

Cons:

- Stores directly identifiable data on a server
- Power imbalance - technology only accessible by big corps and government
- Systemic bias determines enforcement
- Data is determinative

3) Facial Biometrics:

These are secondary features extracted from the face like head movement, eye-movement, heart rate, muscle movement, expressions, assumed emotions, predicted gender, predicted age etc. Like the others, facial biometrics typically uses a webcam to find a face within an image or video, identify the facial features, and track how they move.

You see this capability in everyday devices and photo editing tools that map and move to a face.

Facial biometrics are typically, but not always, conducted with explicit consent from the individual and do not require identity to be useful.

Pros:

- Unlocks contextual information - attention, expression, emotion
- No direct identity needed
- Often used in aggregate
- Data is informative
- Often requires explicit consent

Cons:

- Processing power required means videos are stored on servers

Explanation of Terms

1) Determinative vs Informative Data:

When looking at automating information, the question comes down to the role and influence a human has to play in the final decision.

- **Determinative** - the machine will make a decision without a human verification. Facial verification and Facial Recognition technology fall into this category.
- On the plus side, determinative data can be scaled up and left to run on its own. Conversely, this means any systematic bias will impact many more people than anticipated.
- **Informative** - the machine will inform a human who will make a final decision. This information is rarely binary and needs to be interpreted by an individual before invoking the action.

On the plus side, informative data can rarely be used in isolation from human judgement. Conversely, it has limits to scalable applications and needs to be embedded within a human workflow.

2) Risk & Regulation Considerations:

Like most technology, it depends on the use case and the user, but most of the risk comes down to access and control: access to the technology and control of the decision stemming from the technology.

All of the technologies use artificial intelligence and are subject to systemic bias resulting from training data and access to real-world use during trial periods.

This is the first risk that should have regulation across the board, requiring a reliability test across two factors:

- **Trainability** - the number of use cases required to achieve acceptable accuracy as set out by the use-case experts. Ideally, better than a human practitioner, unless where highly specialised, in which case more scalable than a human practitioner.
- **Transferability** - Ability to apply an algorithm trained in one use case to another with acceptable transfer rates based on similarity of use case.

Now, specific regulation considerations related to the different areas of recognition:

3) Facial Recognition:

Facial Recognition holds the most amount of use cases and the highest risk. At the core, it is used to control access and keep a history of individuals which catalyses a power imbalance between the public and law enforcement agencies. The public has no control or little recompense against a judgement made by facial recognition that can only really be accessed and funded by the state.

Technically, Facial Recognition holds the highest security risk for three main reasons:

1) It is determinative, removing any human ownership
2) It controls access to necessary products and services
3) It cannot be anonymised by its very nature

Socially, Facial Recognition also holds the highest social risk:

1) It can be systematically bias creating unequal law enforcement

Regulation Consideration for Facial Recognition:

- **Watchdog** - For these technologies, regulation will need to account for the security of the hosted database and abuse of power by those who have access to it.
- **Algorithm Auditors** - Suggest implementing AI and cyber security auditors for systematic bias and algorithmic overfitting.
- **Penalties for abuse of power** - also severe penalties towards agencies for misuse.

4) Facial Verification:

Facial Verification is a benign technology. While it does identify an individual, it is only to match it to a government issue ID. Like snapchat, it does not need to track a history, but can still provide tighter security access controls than keys, passwords and even fingerprints.

Regulation Considerations for Facial Verification:

For Facial Verification, the appetite to migrate commercially into Facial Recognition is high.

- **Realtime use only** - restricted use cases in verification defined by matching capabilities that do not reference a stored database.
- **Device only** - restricted to device only use cases

Otherwise, it is Facial Recognition.

5) Facial Biometrics:

Facial Biometrics holds less direct risk because it is informative to a human, meaning the human makes the final decision acting upon all information including Facial Biometrics. Medical use cases can pose a higher risk if positioned as determinative, but this is not recommended. Rather they should inform a doctor to make a final decision.

Facial Biometrics risks vary by type, but can include the following systematic bias:

- **Facial coverings** - beards, glasses, burkas, niqabs may prevent access
- **Dark skin** - low contrast can be difficult for cameras to read and many algorithms do no work well on dark skin - this is improving
- **Low Quality Images** - prevents those with low quality devices to participate
- **Expertise Required** - emotion interpretation often needs a psychology degree

Regulation Considerations for Facial Biometrics:

In addition to the ICO (Information Commissioner's Office), practices already established suggest the following regulation requirements for providers of Facial Biometrics:

- **Certification Qualification** - create a Ground Truth data source with the Turing Institute that is government approved for companies to certify against.
- **Ethical Code** - similar to the ICC/ESOMAR Ethics Code & Guidelines (https://iccwbo.org/publication/iccesomar-international-code-on-market-and-social-research)
- **Explicit Consent** - Clear consent and frequent consent checks for ongoing use cases
- **Training data separation** - meaning that market data cannot be used for training data except where explicitly identified and consented by the participant along with
- **Medical use case approvals**

Andrew Bud CBE, CEO and Founder, iProov

APPG AI Evidence Meeting – Global webinar 6th June 2020

I am Andrew Bud CBE, founder and CEO of London-based iProov, which started up in 2013 and now employs 60 staff worldwide. I am an engineer and serial technology entrepreneur with interest in regulation, and I lead a company built substantially on AI technology. iProov is a supplier of face-based biometrics to sectors including finance, travel, healthcare and immigration. We believe that such systems can be a real force for good in the world, simplifying online life and enabling the digital economy whilst ensuring, through simplicity and security, that it can work for all sections of society.

My evidence addresses the difference between face recognition and face verification, two applications that are often elided and shouldn't be. I will also mention the relevance of an Adequacy Decision by the European Commission.

iProov authenticates online users, whatever their type or brand of device, very simply and securely using face verification. We match a user's face to a trusted image; this might come from their passport or driving licence if they are enrolling, or from a trusted enrolment if they are a returning user. A person asserts an identity online, and we verify their face to check whether that assertion is true or false. Our speciality is to assure, very simply and usably, that the user not only resembles their trusted likeness but is a real human being physically present at the time of authentication. In this way, we protect users against impersonation by photos, videos, replayed recordings or deepfake synthetic spoofs of them.

Our service does not need to know the real identity of the user. Our customers, including the Home Office EU Settlement Scheme, the NHS and many European enterprises, refer to users by anonymous pseudonyms when working with us. Using measures like this, face verification

systems can be designed for privacy.

In protecting citizens, the crucial consideration is rarely a technology itself, but how it is applied. The applications of face matching are radically different. The key distinction is between those applications that are (i) for the benefit of the individual citizen, chosen by the individual citizen and with the consent of the individual citizen and (ii) those that may be for the benefit of the wider society, but in which the user is given neither choice nor benefit from its use. How does this distinction apply to face verification systems such as iProov?

Firstly, our user is informed of what is happening. The application (such as NHS Login) tells the user their face is about to be verified. Then our user interface shows them a stylised image of their face. Pictures are more powerful than words, and the sight of their own features conveys to the user the unmistakeable message that their face is about to be imaged.

Secondly, the user consents. Verification is used to confirm a claim the user makes about who they are. It is not used to identify them. They initiate the process that concludes with the verification, and consent is further assured along the way in the app and user interface. Consent is an absolute legal requirement for normal uses of face verification, according to Article 9 of the GDPR.

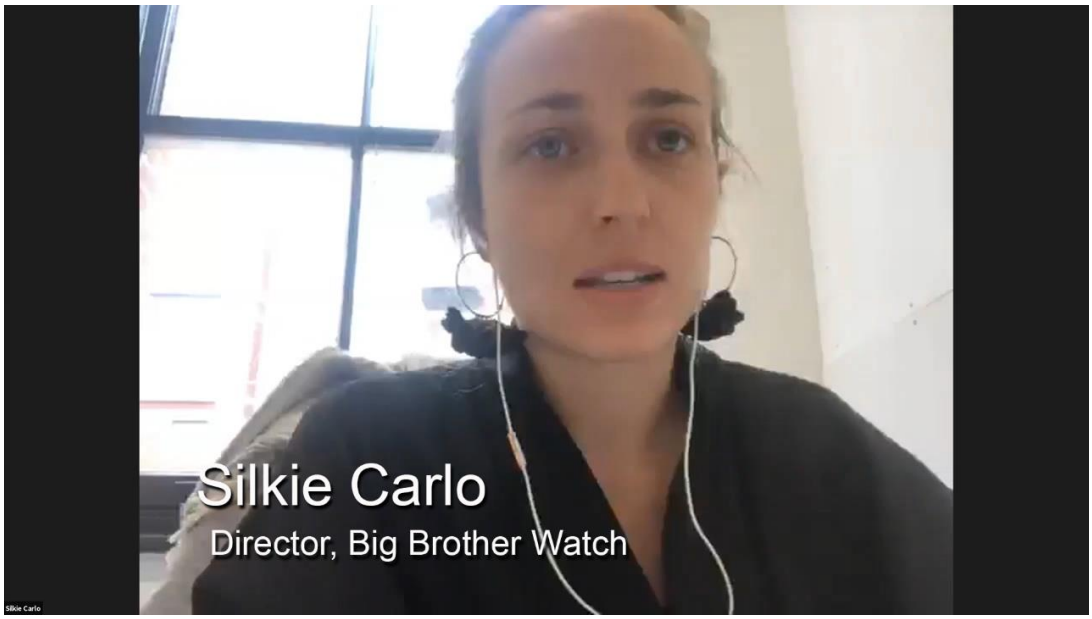
Thirdly, the user directly benefits. Face verification facilitates and secures their access to information and services. It makes the experience easier, more accessible and more inclusive, and yet protects them by securing their identity against impersonation or compromise. In the biometric balance, such user benefit must have its due weight.

There are two recommendations I would like to offer Parliamentarians:

1) Firstly, the difference between face recognition and face verification is fundamental. Regulation of face recognition in surveillance must focus on the threats to privacy and the consequences of misidentification. The risks to privacy and civil liberties from face verification are substantially lower, provided GDPR is complied with in respect of consent, usage and security. The benefits to individual users are considerable; secure online enrolment at home is in practice not possible without it. Regulation that does not acknowledge the profound impacts of knowledge, consent and benefit could deprive users of these benefits for no purpose.

2) Secondly, please consider the impact if a European Commission GDPR Adequacy Decision, or some effective alternative, is not agreed by 31st December. Without such a Decision, it will become illegal for British companies to process the personal data of European citizens without some special legal gymnastics, which will be resisted by major European customers. We have laid contingency plans, but the damage caused to the export potential of the entire UK-based AI industry would be hard to exaggerate.

Silkie Carlo, Director, Big Brother Watch



APPG AI Evidence Meeting – Global webinar 6th June 2020

Big Brother Watch is a privacy and civil liberties organisation. We are non-profit and we have been working on facial recognition for several years now. We have been leading the campaign against the use of live facial recognition in UK in particular. We did the first comprehensive report on its use in the UK in 2018 and we initiated legal action shortly afterwards.

I wanted to first of all acknowledge that I will speak very briefly, and mainly about live facial recognition and of facial recognition as used by corporations and by authorities, although I appreciate there are interesting research and scientific advocations, particularly in the field of emotion recognition. From a civil liberties point of view, of course, the former is more of interest to us. The very prospect of emotion recognition being used for commercial gain is absolutely frightful and unthinkable. I think this is an example of the quite rabid surveillance capitalism that is seeking to profit and scommercialise on every kind of organic and emotional output that humans have, and replacing natural organic processes with technology solutions where they're really not needed, and in ways that often ultimately lead to people being exploited or sometimes misunderstood. I sincerely hope that this is not something that we see more of.

Given the pandemic, there is now the risk of surveillance getting more and more under the skin; when we are looking at an emotion recognition, perhaps even into the mind. I cannot imagine a riskier site of surveillance for civil liberties and human rights. Let us be incredibly cautious about that. Of course, that's why we need regulation. Regulation is what the companies selling this kind of technology want, of course, because it is a life raft for something that is actually very dangerous to populations. What we really are talking about is the search for standards. Where that search begins, perhaps even in some of these technologies where it ends, is with human rights. We look at the rights impacts, and they are extraordinary.

Turning to live facial recognition in particular; first, obviously, it poses a huge risk to privacy and it also poses a risk to freedom of assembly and freedom of association. It turns the presumption of innocence on its head because the use of it - particularly by police - as a random checkpoint in a public place is not sustainable. It will end up on CCTV networks if Parliament does not act now. Individuals are being checked to see if they are criminals. That is not how identity checks have ever worked in this country. We do not have that approach with fingerprints, we do not have it with DNA. We should not have it with facial biometrics. Yet, tens of millions of people have already been scanned in this country, many of whom did not even know about it. That is a breach of the Data Protection Act in and of itself. Well over 3000 people have been misidentified by live facial recognition in this country. Our researchers found the technology being used shopping centres, museums, stadiums, and even protests.

I have spent a lot of today trying to find out if the police were using live facial recognition at the protests yesterday. Which I understand now they were not, but there was so much fear about that among the demonstrators, and everyone's had their photos taken. It may very well be possible that post-event facial recognition is being used. I think that is where the regulatory debate needs to happen. We do not need to completely reinvent the wheel. We already have biometrics, we have fingerprints. We have DNA. There is clearly a use for facial recognition technology in a forensic setting. That needs to be regulated in the way that other methods and approaches are regulated.

I think the question is much more straightforward. In the case of live facial recognition. Again, we will look to pre-existing standards. We do not have identity checkpoints in this country. We do not ask millions of people at one time to go undergo an identity check during everyday life.

Finally, I just, I want reference what we talk about when we talk about misidentifications and the general way that live facial recognition works. I would encourage everyone to think about this outside of simply the relationship between an algorithm and an individual. What I have seen on countless observations, is that it actually feeds into a culture of over policing and suspicion, where you have lots of people suddenly being asked to account for themselves in a way that you would not ordinarily see the police do.

On one occasion, I saw police using this in London. They misidentified and stopped a 14-year-old black boy, which is commonly what we see when this technology is used. He was pulled over by four plainclothes police officers who dragged him across the side of the street. He was absolutely terrified. They demanded his fingerprints and did an identity check. They wanted to see his ID card. He does not have an ID card; he is a child in his school uniform. I think that in those moments he formed a very, very negative view of the police - as the police as a very oppressive institution. That would be a fact-based analysis of what he has just experienced, so we cannot sminimise it. A lot of people think that as soon as you just identify yourself as an innocent person, it is all said and done. That is not the case, this has a long-lasting impact.

My ultimate and final messages is that I think the case to ban live facial recognition is overwhelming and that should be a priority. Then the other areas of biometric technology need to be looked at for regulation.

Contact

APPG AI Secretariat

Big Innovation Centre
62 Wilson Street
London EC2A 2BU
United Kingdom

info@biginnovationcentre.com
www.biginnovationcentre.com

appg@biginnovationcentre.com
www.appg-ai.org