



# Deepfakes: a human challenge

How can tooling be used to assist humans in deepfake detection?



**Bringing Ingenuity to Life.**  
[paconsulting.com](http://paconsulting.com)

# Foreword

A decade ago, the advent of the ‘digital age’ and internet-enabled crime was one of policing’s most pressing priorities.

A decade ago, with a team from PA Consulting, I supported the development of the framework of digital capabilities to enable policing to understand and respond to the seismic shift from ‘traditional’ to ‘online’ crime.

In that new digital world, it was clear that the effectiveness of policing’s response would be determined by their ability to adapt to technological change and rapidly develop new solutions.

The intervening years have seen huge effort and investment aimed at transforming the tanker into an agile fleet that now responds to a volume and scale of ‘internet-facilitated crime’ and cyber-crime in ways that were previously unimaginable.

Adaptability, modular architecture and scalability have all been watchwords for these digital transformations. But now is time to test the fleet’s response to a new emerging technological challenge – artificial intelligence and deepfakes.

Deepfakes will again change the relationship of humans with the digital world. There is no technological silver bullet to mitigate the impact of deepfakes on crime.

We must continue to respond with adaptability, innovation and ingenuity to the rise of deepfakes, and the Home Office’s Deepfake Challenge is the first step in engaging the public and private sectors to do just that.

At PA, we believe in the power of ingenuity to build a positive human future, and we believe that together we can tackle this challenge.

**Ottoline Warner**

**Public Sector Innovation & Digital  
Partner – PA Consulting**



# Contents

---

01 Introduction

---

02 When seeing is no longer believing

---

03 Digital media evidence in police investigations

---

04 Testing digital media evidence in courts

---

05 Deepfake CSAM: A worrying new trend

---

06 Empowering the public through education

---

07 Conclusion

---

A Appendix

# Executive Summary

*This report is in response to the Home Office's Deepfake Challenge - June 2024, addressing the question:*

*How can tooling be used to assist humans to detect deepfakes?*

**The rise of deepfakes that are indistinguishable from 'real' videos, images or audio will fundamentally change how we perceive and use digital media online.**

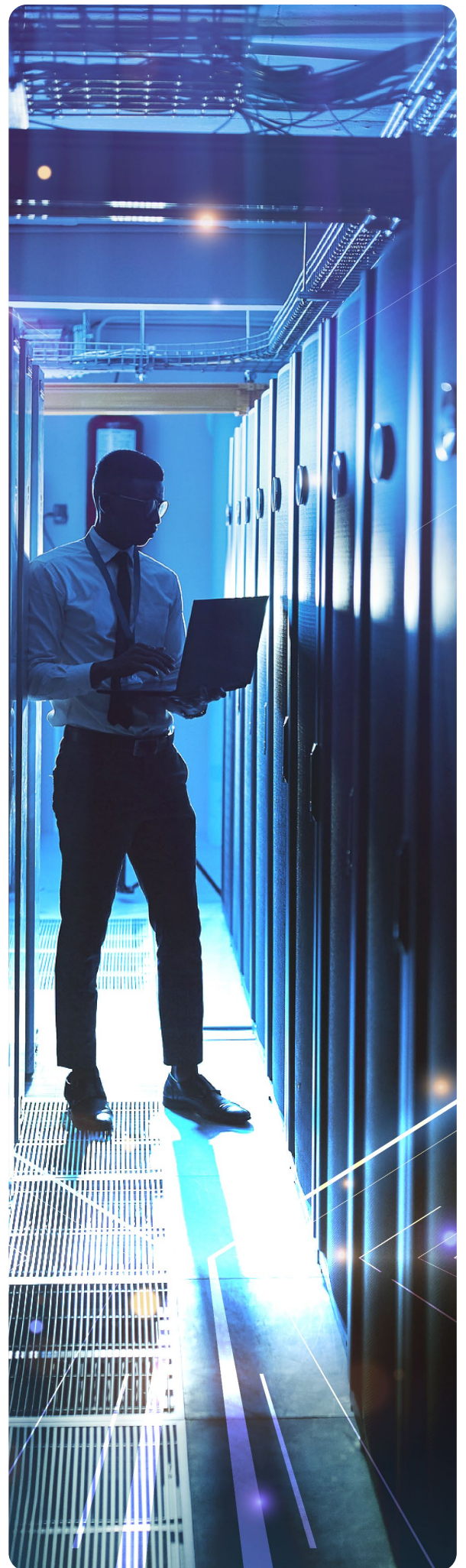
To understand how we can use tooling to help us detect deepfakes, we must consider the wider operating models in which we need to make detection decisions. Tooling will never provide a single definitive ruling on the veracity of digital media because as fast as detection tooling advances, so too do the techniques used for deepfake creation.

As detection moves beyond human capability, tooling will still be critical, but must be coupled with operating model change and a systematic approach to incorporating detection technology.

This will include:

- Strengthening existing trusted sources of evidence for policing through hashing technology and cryptography to maintain trust in digital evidence chains
- Updating processes for evidential assessment and equipping officers with the tools and skills to make informed investigative decisions on deepfakes
- Preparing courts with the roles and skills needed to test digital evidence and establish new case law
- Equipping officers and analysts with the triage processes, tooling and knowledge to tackle deepfake Child Sexual Abuse Material (CSAM)
- Engaging with the ecosystem of organisations that tackle CSAM to establish shared taxonomy and systematic approach to categorisation and response
- Helping citizens to protect and inform themselves through tailored public education campaigns and tools

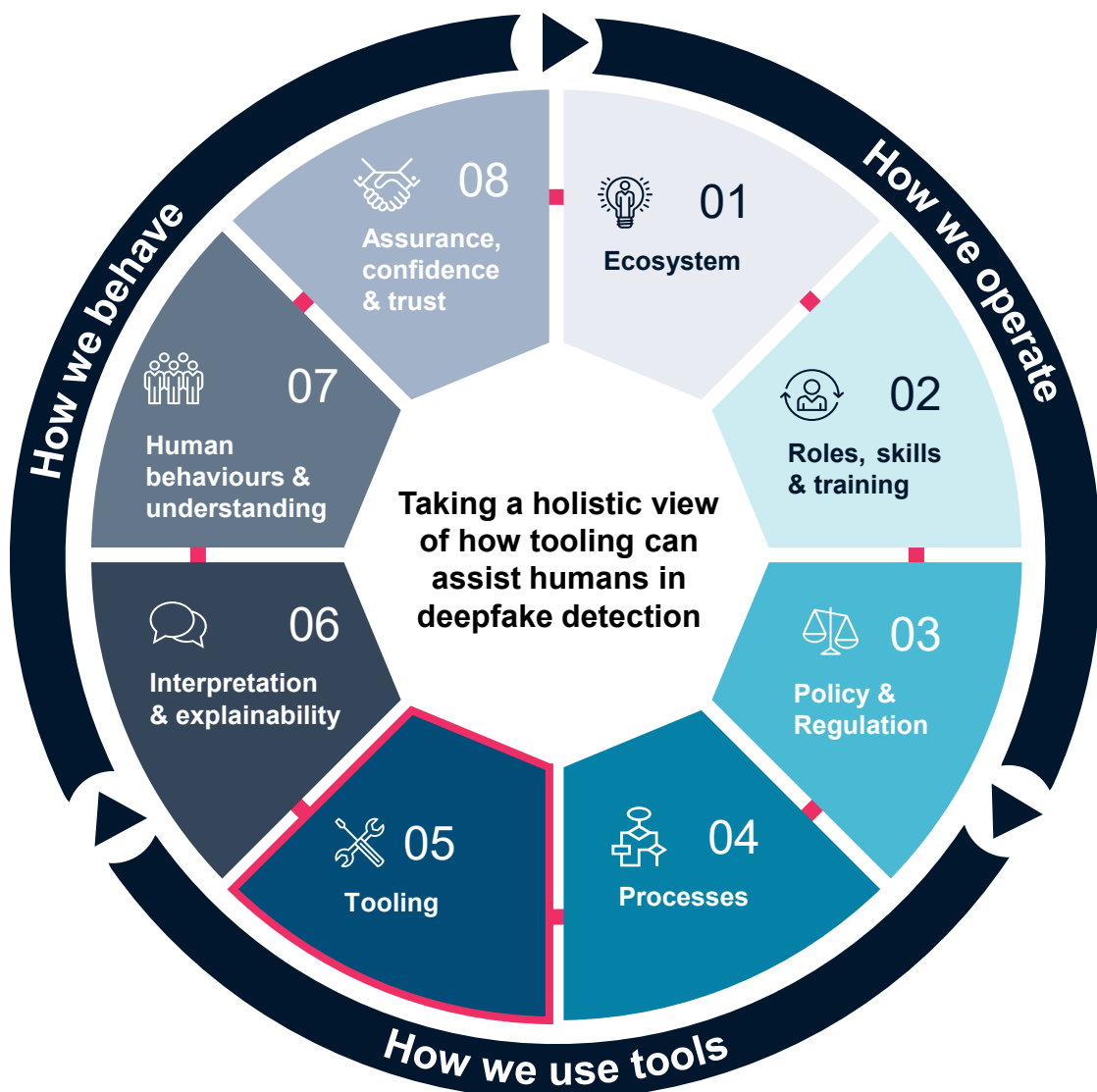
Each of the case studies that we have considered demonstrates how tooling can be effectively combined with a systematic approach to change, collaboration across organisations and ecosystems and an understanding of human behaviour to assist humans in detecting deepfakes.



# Through a holistic view we can understand how tooling could assist humans in deepfake detection

**As the increasing presence of deepfakes disrupts the way we interpret the world around us through digital media, we need to equip ourselves, our law enforcement and the wider justice system with the ability to make informed decisions.**

Whilst tooling will be needed to help detect deepfakes, it will also be necessary to consider wider operating models and ecosystems to fully equip humans to use that tooling to make decisions effectively.



**We believe that this approach is fundamental to finding coherent and effective solutions to a systemic problem.**

# Deepfakes will impact ways of working and change the nature of threats

To understand how humans will use tooling we must consider use cases from both areas of impact.

## Deepfake-impacted procedures



Threats which may not fundamentally change but **processes and investigation will be impacted** by deepfakes. I.e:

In which digital evidence may be used to pervert the course of justice, for example in volume crime and serious crimes such as:

- Neighborhood harassment
- Domestic abuse
- Serious sexual offences
- Kidnap
- Murder

Increased resourcing and effort required in investigations, as well as potential for additional increased harm

Public trust and confidence is undermined through the presence of deepfakes which will impact trust in policing as well as courts procedures

Increased resourcing and effort required to provide clarity and assurance

## Deepfake-enabled crime / threats such as Threat areas that are changing because of deepfakes. I.e:



### Intimate image abuse

New avenues of abuse / harm through creating and sharing sexually explicit deepfake imagery of adult victims

### CSAM

New avenues of abuse / harm and challenges with changing offender behaviours that causes increased harm and re-victimisation

Increased direct harm to victims through significantly increased scale of crime, as well as increased re-victimisation

### Fraud

Deepfake tooling offers new methods to conduct fraud more effectively increasing the scale and potential severity of the crime

Increased direct harm to victims through increased scale and 'effectiveness' of the crime

### Mis / Dis information

Individuals, groups and states are already using deepfake technology to spread false information more convincingly at scale

Undermines society's ability to trust digital media and information

Where crimes become deepfake enabled, they pose a significant increase in the volume and harm caused to victims. Whereas procedures impacted by deepfakes will significantly increase the effort and resources required, which may indirectly increase harm.

# To fully understand these impacts we consider case studies that address different aspects of the challenge

The case studies we have selected look at a range of challenges and issues to give us insight into both specific threat areas and to draw wider conclusions on how best humans can use tooling in the future to assist with deepfake detection.



## Deepfake-impacted procedures

### Digital Media Evidence in Police Investigations

The use of digital media as evidence is heavily used in policing across all crime types. We consider the sources of and processes that support the use of digital media as evidence to identify how deepfakes will impact existing operating models, and how it will need to look in the future.



## Deepfake-enabled threats

### Deepfake CSAM: a worrying new trend

Deepfakes are changing offending behaviour and increasing the volume of CSAM. Working with the Internet Watch Foundation and Childlight, we explore how this will impact ways of working and what must be done to get ahead of the challenge.



### Testing Digital Media Evidence in Courts

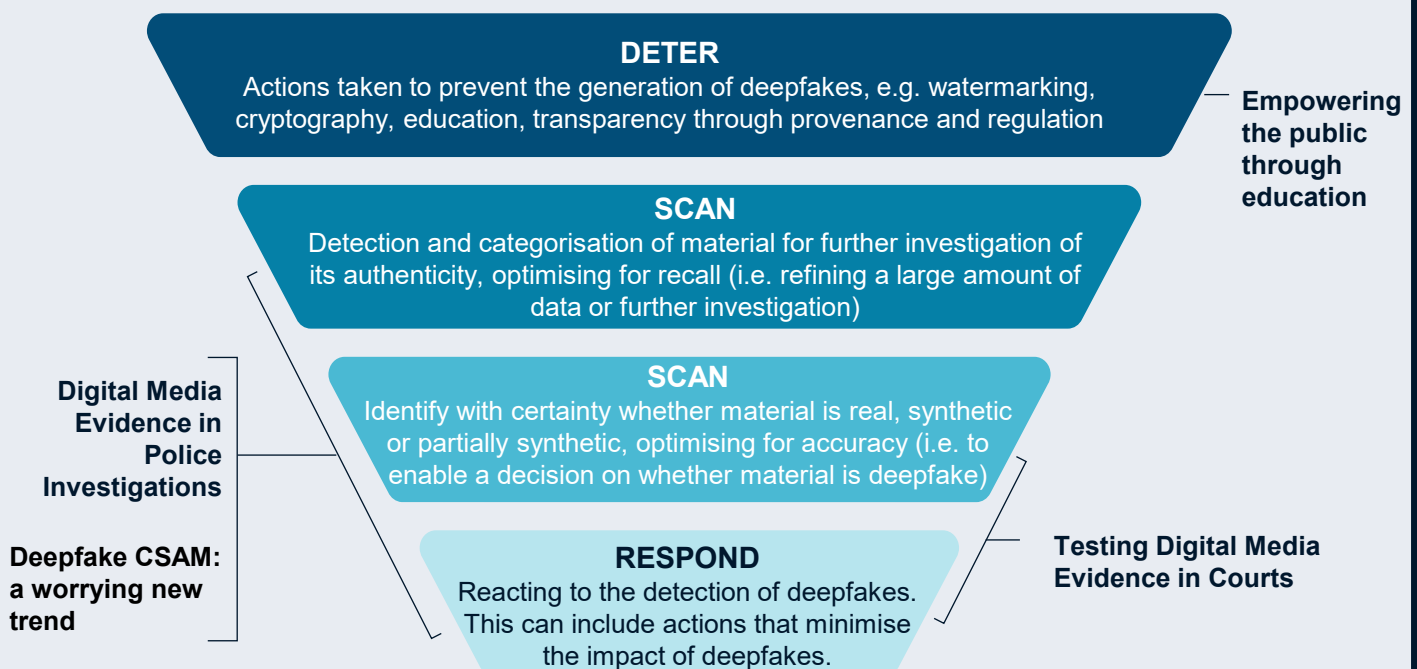
Digital media as evidence must eventually stand up in a court of law to be effective in the pursuit of justice. We explore how initial evaluation and assessment will be made in the future and how deepfakes will disrupt courts in the near future.



### Empowering the public through education

The role of public education is critical in helping citizens make informed decisions on what is real across many areas of digital engagement. We imagine how we can help to engage and educate the public in educating themselves.

These case studies also address different areas of the deepfake detection lifecycle:



# Complex problems need a collaborative approach

We have brought together experts from across PA sectors and capabilities, specialist industry partners alongside clients and organisations working at the forefront of these challenges to provide insight and expertise.

<p><b>Policing &amp; Justice teams</b> </p> <p>PA works extensively with justice agencies - law enforcement, prosecutors, courts, prisons and probation - in the UK and internationally. We have often supported the adoption of new technologies to improve operational processes, ranging from the introduction of ANPR, to the roll-out of alcohol monitoring tags, and the adoption of AI in the courts.</p>	<p><b>Online Safety team</b> </p> <p>PA has been working for ten years with UK and international governments, law enforcement agencies, the technology industry and third sector partners to research how online harms, and the response to them, are evolving globally (considering offender and victim behaviours, technology developments and socio-economic factors).</p>	<p><b>Digital UX / UI</b> </p> <p>PA focuses on creating innovative and user-centric UX / UI designs that enhance user experiences and drive commercial success. Our approach involves a deep understanding of user needs, combining creativity with analytical insights to craft intuitive and engaging digital products and services.</p>
<p><b>AI capability</b> </p> <p>We develop and implement AI strategies and AI OpModel, we embed AI models into custom hardware. Our teams are well-versed in AI policy, regulation, and ethics, guiding the responsible application of AI. With expertise in scaling AI and achieving Enterprise AI, PA's breadth of knowledge and experience uniquely positions us in the market.</p>	<p><b>Human Insight</b> </p> <p>We have a team of 70+ research professionals, all experts in a range of quantitative and qualitative research methods both traditional and innovative, working across the private and public sector. We use mixed-method insight and strategic thinking to understand people, businesses and organisations and represent them, driving better design, innovation, strategy and experience.</p>	<p><b>Innovation team – Ingenuity festival</b> </p> <p>We took this challenge question to one of PA's Ingenuity Festivals. These day-long events bring together our clients' challenges with multidisciplinary innovation teams from PA. Following a structured innovation approach, team generate ideas, challenge assumptions and develop solutions over the course of the day.</p>
<p><b>Advai</b> </p> <p>Advai's platform specialises in stress testing AI and ML models, providing real-time metrics to identify potential failures. It tests datasets and models pre &amp; post deployment, enabling anomaly detection and performance optimisation. This approach helps developers enhance the reliability and stability of AI systems across various applications, minimising risks and ensuring robust performance. They have developed a system model for what a generic deepfake system would look like and plan to demonstrate their assurance tools against deepfake detection tools.</p>	<p><b>Videntifier</b> </p> <p>Videntifier provide Videntifier Nexus, a platform for identification of known illegal visual content, using hash databases from various sources. Combining identification technology, by fingerprinting media, with hash databases, it is a video and image matching solution – this supports the monitoring, detecting and response to illegal digital content, and it is being used to support a European law enforcement agency, speeding up case processing by automatically detecting known CSAM images / videos, and identifying duplicates.</p>	<p><b>T3K</b> </p> <p>Provide AI solutions for finding and classifying illegal and harmful content. One of their solutions, CORE, is an image and video classifier that screens data to find illegal and inappropriate content such as CSAM (focusing on first generation), terrorism and more, and classify it. This API-only solution can be integrated into existing content review workflows and is the world leader in child abuse detection solutions.</p>

With thanks to the following for their time and input:





# 02

---

When seeing is  
no longer believing



# Society is approaching a crisis of trust in digital media

Since the advent of photography, images and videos have been a reliable proxy for 'reality'. Historically, the resources and effort needed to fake and share images was generally prohibitive until the advent of the internet and digital media.

Today, the advances in artificial intelligence technology means that the ability to 'deepfake' digital media, such as video, images and audio, is fast becoming so sophisticated that these media are indistinguishable from reality to the naked eye and increasingly to forensic tools and machines.

Humans are naturally inclined to believe what they see; deepfakes leverage this trust, creating imagery so close to reality that it captures the essence of believability.

The cognitive biases that govern our perception are expertly played upon, making the task of discerning real from fabricated a formidable task for individuals without forensic tools.

In particular, audio and video are viewed as having more 'testimonial authority' than other representations of the world.

The contextual cues that often guide our judgment are meticulously replicated in deepfakes, leaving little room for doubt. The interplay of light, the gravity of expressions, and the coherence of movements are all orchestrated with a precision that mirrors reality, blurring the lines between what is real and what is artificially constructed.

From a non-technical perspective, the difficulty in detecting deepfakes stems from their exploitation of trust and the expectation of truth in visual media. This is predicted to have far-reaching impact on crime, victims of crime and how law enforcement and government needs to respond. [1]

If seeing, now, is no longer believing, deepfakes will have fundamentally changed our relationship with digital media, causing us to call into question the veracity of all the information we see and hear online.

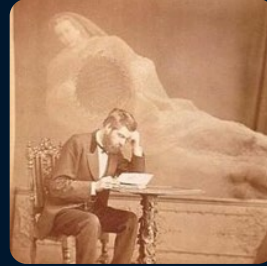
In the short term, as society seeks to understand this paradigm shift, we are likely to be easily fooled, but as a recent study demonstrated, as people are exposed and aware of deepfakes, they become overall less trusting in audio and video media types in general as well as in their own ability to discern a manipulated media [2].

1. Europol (2022), Facing reality? Law enforcement and the challenge of deepfakes, an observatory report from the Europol Innovation Lab, Publications Office of the European Union, Luxembourg.
2. Weikmann, T., Greber, H., & Nikolaou, A. (2024). After Deception: How Falling for a Deepfake Affects the Way We See, Hear, and Experience Media. *The International Journal of Press/Politics*, 0(0). <https://doi.org/10.1177/19401612241233539>
3. <https://en.wikipedia.org/>

## 'Deepfakes' – a definition

"Deepfake, synthetic media, including images, videos, and audio, generated by artificial intelligence (AI) technology that portray something that does not exist in reality or events that have never occurred"

Britannica



*William Mumler's Spirit Photography*  
[3]. 1870



*The Cottingley Fairies*  
1917



*Colin Evans, the Levitating Man*  
[3]. 1937



*Pope in a puffer jacket*  
[3]. 2023

**We must equip ourselves with the understanding, tools and techniques to navigate this new paradigm**

# As deepfake tooling and techniques advance, it becomes a human and a technical challenge



## Technology sophistication

Deepfake technology presents a formidable challenge in detection due to the level of sophistication necessary to overcome human perception and the nuances of authenticity.

At the heart of the issue lies the sophistication of the algorithms that generate deepfakes.

These algorithms are adept at learning and mimicking the subtlest of human expressions and movements, creating images and videos that are indistinguishable from genuine content to the untrained eye.



## Coupled with commoditisation and volume

The challenge is further compounded by the democratisation of technology. As the tools to create deepfakes become more accessible, the pool of potential creators widens, leading to a proliferation of content that ranges from the benign to the malicious.

This vast array of generated content overwhelms the capacity for manual verification, necessitating the development of automated systems that can keep pace with the rapid generation of deepfakes.



## Create a significant challenge

The difficulty in detecting deepfakes is a matter of technological advancement and a reflection of our own vulnerabilities in perception, the erosion of trust in media, and the sheer volume of content that demands scrutiny.

It is a multifaceted problem that requires a nuanced understanding of both human psychology and the evolving landscape of artificial media.



# To keep up with the technology arms race, we must consider the full range of tooling available...

## Digital forensics

**Digital forensics** involves a suite of statistical and conventional image processing techniques, which use various analytical approaches to examine the authenticity of digital media. These include:

- **Statistical Analysis:** Examining the statistical properties of an image or video, such as pixel correlations, to detect anomalies that may suggest manipulation.
- **Image Processing Techniques:** Utilising conventional methods like error level analysis, reverse image search, and frequency analysis to identify signs of tampering.



**Strengths:** Non-Reliance on Machine Learning means these methods are typically less data-dependent and faster to apply in some cases. A broad range of tools are available, each with its own method of uncovering evidence of fakery.



**Weaknesses:** Sophisticated deepfake methods may evade detection by these non-ML techniques, especially as technology advances. Due to the need to use digital forensics, specialised knowledge and skills may not be accessible to all users.

## Metadata analysis

Metadata analysis examines the **digital footprint and auxiliary information of the content**, such as timestamps and editing history, embedded in digital media. Data points not directly related to the content of the files themselves. **Examination can reveal traces of manipulation.**



**Strengths:** It is a **non-intrusive** method that does not rely on the content itself.



**Weaknesses:** **Metadata can be easily altered or stripped, and not all deepfake generation processes leave behind detectable metadata traces.** This limits the applicability of metadata analysis in many cases.

## Biological Signals Analysis

BSA techniques analyse biological cues (e.g., blinking patterns, heart rate, subtle biological movements) to detect inconsistencies in deepfake videos. These cues provide indirect evidence of manipulation.



**Strengths:** BSA focuses on **involuntary human traits**, such as **eye blinking or heart rate**, which are difficult to replicate in deepfakes. Where image data is perceived/claimed to be of sufficient resolution movement of blood within arteries can also be used. This can make **BSA a robust approach** against deepfakes that do not account for these biological signals.



**Weaknesses:** BSA may be **limited by the quality and resolution of the video**, as detecting subtle biological signals requires high-fidelity data. Additionally, **advanced deepfakes may learn to simulate these signals**, reducing the effectiveness of BSA.

# ...and how they can be combined for best effect

## Contextual Analysis

Examining activities such as **when and by whom an image was first posted, and how it is distributed** before becoming 'mainstream' (e.g. reputable news outlets or anonymous posters on hyperbole social media sites). **Subtle image content can also be examined compared against reality**; from sophisticated to banal, cues in the image are sought reveal inconsistencies (e.g., star patterns present the wrong location/time of year).



**Strengths:** Contextual analysis looks beyond the content to consider the broader context, such as the source and the surrounding narrative. This can help identify deepfakes that are part of misinformation campaigns.



**Weaknesses:** : This method requires a **comprehensive understanding of the context** and may not be applicable in situations where the context is unknown or irrelevant. It also **relies on external information** that may not always be available or verifiable.

## Deep Learning

Deep learning models, such as Convolutional Neural Networks (CNNs), dominate deepfake detection. CNNs learn hierarchical features directly from data, making them effective for identifying manipulated content. CCNs aim to identify visual or audio artifacts introduced during deepfake generation. Detecting anomalies, compression artifacts, or inconsistencies helps flag manipulated content.



**Strengths:** Deep learning models, particularly Convolutional Neural Networks (CNNs), **are highly effective at feature extraction and can learn complex patterns** in data. They are **currently the dominant method for deepfake detection** due to their performance on large and complex datasets.



**Weaknesses:** These models require **large amounts of labelled data and significant computational power**. They can also be **less interpretable** and may suffer from overfitting the model to Is not limited to the contextually important information in the image regularisation.

## Neuro-symbolic approaches

Neuro-symbolic approaches combine the strengths of **deep learning and symbolic artificial intelligence**. The deep learning component **utilises neural networks** to process and learn from large training data sets, identifying complex patterns and anomalies that may indicate a deepfake. The Symbolic AI **component applies logic and rules to interpret structured information**, assessing the content at a conceptual level to detect inconsistencies.



**Strengths:** By understanding **both pixel-level details and higher-level concepts**, this approaches can detect deepfakes through **context and reasoning**, not just visual cues. These methods can **recognise logical inconsistencies** or unexpected behaviour in the content, which pure pattern recognition might miss.



**Weaknesses:** The combination of deep learning with symbolic reasoning can be **challenging, requiring sophisticated** model design and training. In addition, effective training of these systems often **demand large, well-annotated datasets**, which may not be readily available

# Deepfakes are not just a technical problem and must be resolved with a holistic approach

Key aspects that must be considered are:

01

## Complexity of Deepfakes:

The sophistication of Deepfake technology has necessitated a collection of diverse detection techniques. This complexity arises, in part, from the use of advanced deep learning algorithms that can generate highly realistic synthetic media.

02

## Continuous Evolution:

As Deepfake technology evolves, so must the detection methods. Ongoing technological evolution suggests that no single method will be sufficient to detect all Deepfakes, as creators continually refine their techniques to avoid detection.

03

## Importance of Public Awareness:

The variety of detection methods underscores the need for public education on the matter. By understanding the common signs of Deepfakes, individuals can become more critical consumers of media and less susceptible to misinformation.

04

## Technological Arms Race:

The wide variety of detection methods reflects 'Action-reaction' dynamics between Deepfake creators and those developing tools to identify them. As new detection techniques are developed, new and improved creation capabilities are created to overcome previous flaws that detection exploited.

05

## Need for a Multi-Faceted Approach:

Relying on a single detection method could be ineffective due to the diverse nature of Deepfakes. A combination of techniques, including deep learning-based, traditional machine learning-based, artifacts analysis-based, and biological signal-based methods, is necessary for wide scale, robust detection.

06

## Accessibility of Tools:

The fact that there are multiple ways to detect Deepfakes also indicates that the tools to create and detect them are becoming more accessible. This democratisation of technology has both positive and negative implications, as it allows for widespread creative use but also facilitates the spread of disinformation.

07



## Confidence in tooling is critical:

Any tools to detect AI generated content are only as effective as the trust and confidence people place in these tools. This is critical when the tools are derived from AI technologies and any output needs to be defended as evidence; validating and assuring their output is as essential as the output itself. There will be a need for technical governance and assurance to ensure that tools are both well understood and trusted.

# 03

## Digital media evidence in police investigations

A case study in deepfake-impacted procedures



# Deepfakes will undermine the veracity of digital media and its use as evidence in policing

**Gathering digital media from police-verified sources and the public domain is an essential part of criminal investigations and prosecutions. 90% of criminal cases are estimated to rely on digital evidence.**

A rise in the creation of realistic deepfake material poses a risk that the use of digital media in policing investigations will be undermined.

This presents a risk that the course of justice could be perverted because of a misleading deepfake, regardless of the source's intention.

In addition, the possibility of undetected deepfakes being used as evidence in police investigations will provide reasonable grounds to doubt the veracity of evidential digital media as a whole.

This will put digital media under a higher degree of scrutiny and erode confidence in its use as evidence and intelligence.

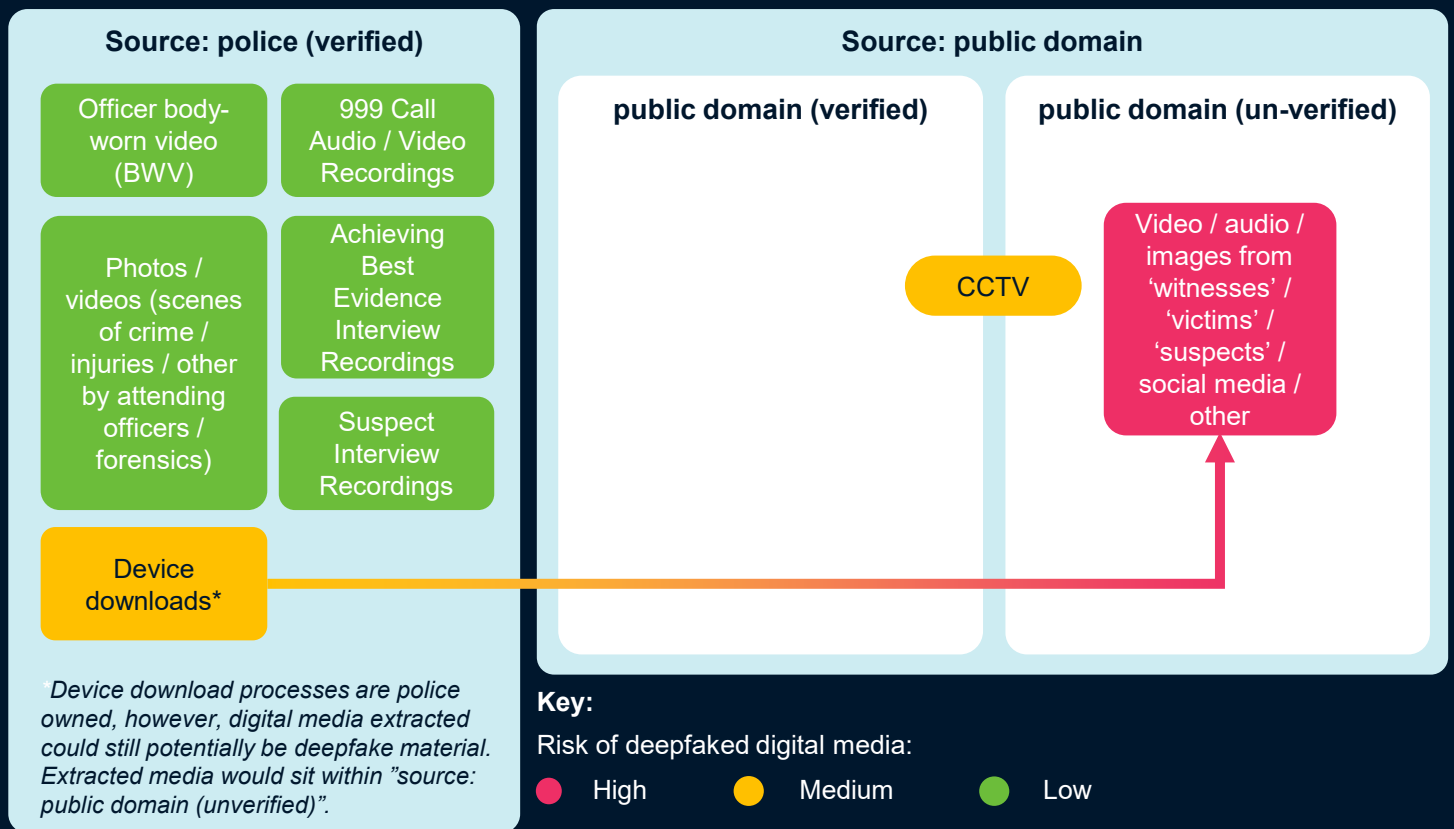
## Deepfake-impacted procedures: policing



Digital media for evidence that will be impacted are:

- **Use of unverified sources from the public domain such as social media** or digital device downloads
- **Use of police verified sources** such as police body worn video (BWV)

## Sources of evidential digital media



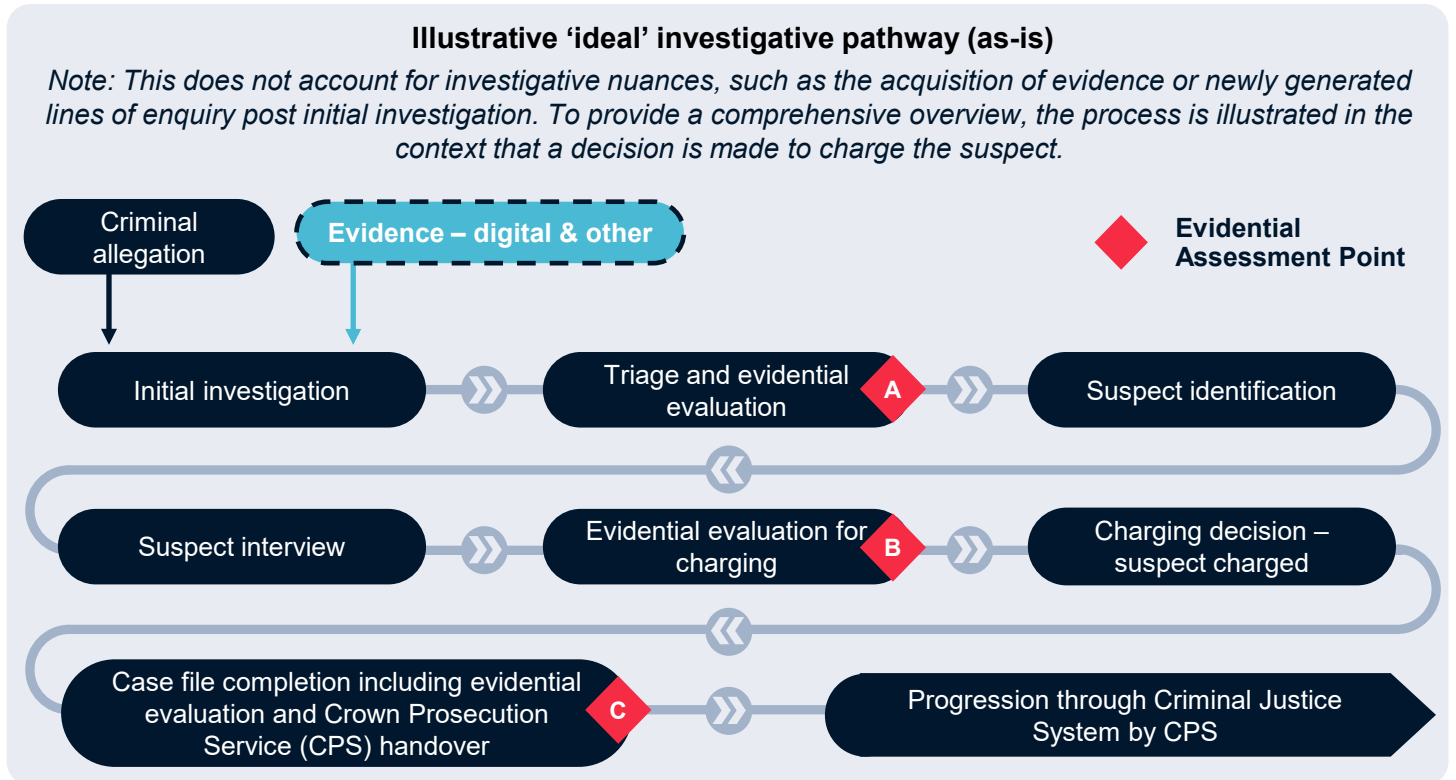
**Policing will need to embed new tooling and ways of working into operational and investigative procedures to assure the veracity of digital media from police sources and the public domain to maintain trust in the use of digital media as evidence.**



# Evidential assessment procedures will need to adapt to manage unverified digital media

**There are three key evidential assessment points in an ideal police investigation where digital media will need to be further scrutinised.**

At these points, the introduction of deepfakes within digital media could undermine decision making and impact an investigation.



**A**

### **Triage and evidential evaluation**

Officers obtain evidence, following an allegation, on scene or e.g an assigned online report.

An **initial assessment of evidence contributes to decisions on next steps**, e.g. making an arrest.

This could include **significant volumes of digital media** that may require a triage process.

**B**

### **Evidential evaluation for charging**

Once a suspect has been identified and interviewed, the officer in charge will complete initial case file preparations to submit to a supervisor for **evidential review and authorisation of a charging decision**.

**C**

### **Case file completion**

**Evidence will be scrutinised through prosecution processes managed by the CPS.**

The officer in charge of the investigation will be responsible for **accommodating any requests made by the CPS in relation to the assessment / provision of evidence**.

**Each of these points will require officers to have the tooling, training and processes in place to factor the presence of deepfakes in digital media into their decision making.**

- For use of digital media from unverified public sources, additional scrutiny and assessment of digital media will be required;
- For use of verified police sources, additional technology solutions and assurance processes to build confidence into the digital evidential chain.

# Officers will need the training, skills and tooling to detect deepfakes

**Deploying tooling will assist investigators in assessing the veracity of digital media and detect deepfakes.**

This will apply to digital media that is at risk of being deepfaked – predominantly from unverified sources in the public domain.

However, to make effective evaluation and decisions it will require policing to adapt its operating model at the key evidential assessment points.

We have considered the following potential adaptations, in reference to this illustrative investigative pathway:

- A** Tooling and training of officers to support fast, early decision making alongside the upskilling and availability of Digital Media Investigators (DMIs) to provide guidance on usage and outputs and how it may impact their working strategies.
- B** Ability to submit digital exhibits to DMIs to provide more in-depth digital media veracity assessments, relying on a more extensive tooling suite, ahead of suspect charging decisions.
- C** Submission of digital exhibits to a dedicated role within digital forensics for comprehensive digital media veracity assessments, supported by use of a complete tooling suite, post-suspect charging.



## Police Constable Olivia

can make quick decisions and assessments of a potential deepfake on scene using her mobile device to run basic detection tooling



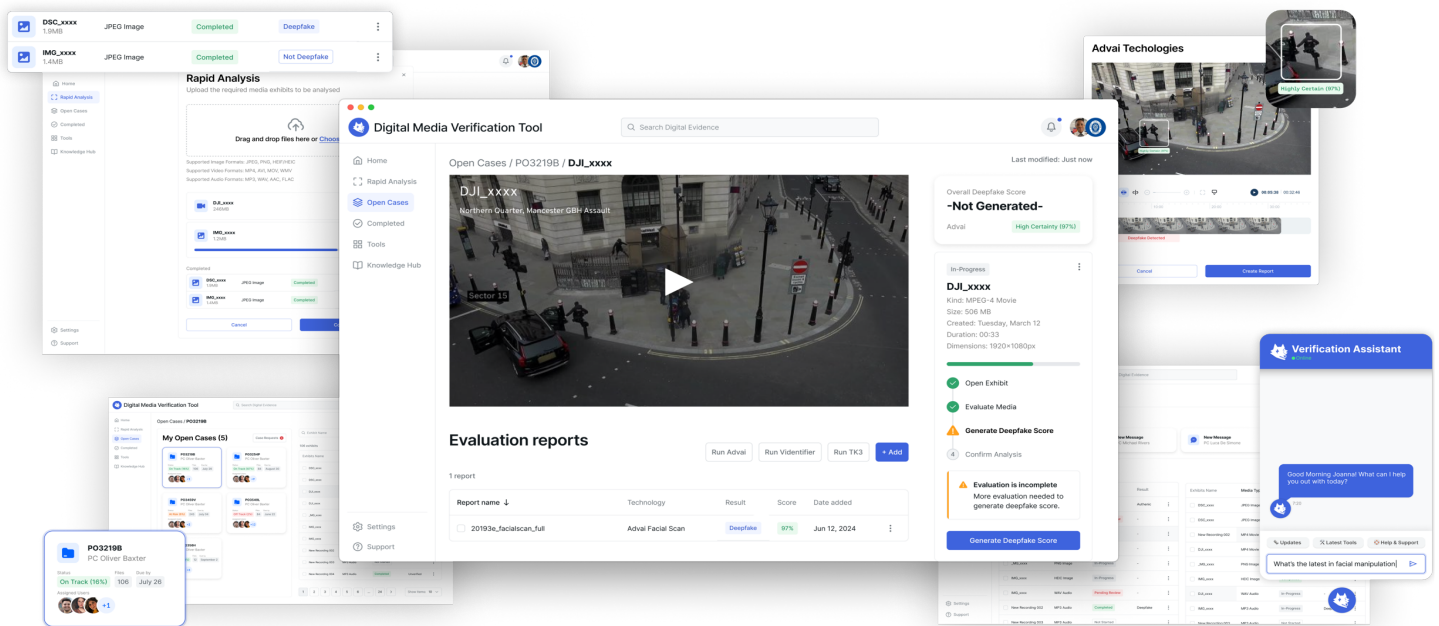
## Digital Media Investigator Joanna

needs to be up to date on the latest in her tooling suite and in deepfake creation techniques to advise investigations and run more in-depth analysis



## Digital Forensics Expert Michael

has an advanced range of deepfake knowledge and analysis tools at his disposal to provide expert insight in complex investigations



**In the Appendices we have developed user personas, scenarios and wireframes to explore what the future needs of policing will be to adapt to deepfake digital media.**



# To assess **unverified** sources, new processes, training and tooling are required



## Processes

*New processes for assessment and referral for deepfake checks*

- Defined processes for conducting and recording deepfake assessments at each stage in an investigation.
- Processes for when / how to pursue accessing guidance via DMIs or Digital Forensics deepfake experts
- Processes coupled with Service Level Agreements to ensure referred assessments are returned in line with investigation progression requirements (e.g. the PACE clock, or CPS action requests).



## Tooling

*Levels of tooling will need to adapt to role and decision-making context*

- Tooling / tooling suites will need to be defined and tested for use by different roles for specific requirements – i.e. 'rapid' tooling for initial on scene assessment vs advanced tooling suite for Digital Forensics experts.
- Integration with case management / tasking tooling to manage referrals to DMIs / Digital Forensics for deepfake assessment requests.
- Where large volumes of media need to be assessed, a triage pipeline to automate assessments within specified parameters.



## Roles, skills & training

*Existing roles will need new training on awareness and impact of deepfakes*

- All training will need to be regularly updated and refreshed to keep pace with changing technology. Rate of refresh will depend on the role requirements.

### **Officers**

- Awareness training for deepfakes and the potential falsity of digital media provided by unverified sources. Understanding that this must be considered when developing working strategies and considering corroborative evidence.
- Training on new processes and tooling to assess the veracity of digital media or submit it to DMIs / digital forensics.

### **Digital media investigators (DMIs)**

- Ability to provide guidance to front line officers supporting uncertainty around the veracity of digital media.
- Ability to use tooling to conduct assessments at a more advanced level, and to interpret and communicate the results

### **Digital forensics (DF) – digital media veracity assessment role**

- Ongoing training on tooling to support them in making advanced deepfake assessments, understanding the mechanisms of tools so they can select the most appropriate for the specific media and requirement.
- Skills and training to supplement automated tool-generated analyses with qualitative analysis and personal evaluation.



## Interpretation & explainability

*All roles will need to interpret tooling results to support decision making*

- Deepfake tooling that uses AI presents challenges to explainability – more advanced roles will need to interpret and explain in different contexts, for example the Digital Forensics expert acting as an expert witnesses in court to explain determinations.
- DMIs and DF experts will need to have the technical abilities to interpret, assess and justify different tools usage when coming to an opinion.

**Critically, the scope of digital media assessments will need to be defined and levels of risk appetite agreed. This will need to depend on the organisation, team and potentially specific case parameters.**

# There are opportunities to assure **trusted** digital sources with technical solutions

**Source: police (verified)**

Officer body-worn video (BWV)	999 Call Audio / Video Recordings
Photos / videos (scenes of crime / injuries / other by attending officers / forensics)	Achieving Best Evidence Interview Recordings
	Suspect Interview Recordings
Device downloads*	

*\*Device download processes are police owned, however, digital media extracted could still potentially be deepfake material. Extracted media would sit within "source: public domain (unverified)".*

**Where digital media are produced with controlled devices the opportunity exists to assure their outputs. This provides certainty around their verification and removes any burden associated with proving their veracity.**

Policing should initially explore how this could be achieved for sources that run the risk of being 'deepfaked' in the public sphere, for example:

- Body Worn Video:
- 999 calls
- Publicly owned CCTV footage
- Footage from products – i.e. Ring doorbells

Approaches to assuring these chains of digital evidence for policing may not prevent the creation of deepfake footage that intends to target or implicate police officers, but it can create more trusted evidence chains used in investigations, intelligence and into courts.

**Introducing a digital assurance will require systematic changes affecting the tooling used to process digital media, the processes used to produce, handle and create it and education to as to the assurance and confidence it produces in its intended audience (e.g. courts, the public)**



## Tooling

*'Hashing' provides a method of assuring police digital sources*

**'Hashing'** is a cryptographic method for creating a unique **digital fingerprint** of digital media. Changes to source content will result in a different digital fingerprint. This enables content to be verified as 'original'. Hashing capabilities added to controlled devices and their supporting infrastructure will provide verification of their outputs' veracity.



## Processes

*Processes will need to ensure they maintain the 'hash'*

Good enabling technology is invisible to the user, but processes for the creation, handling and curation of digital media will need updating to ensure the media is not affected in such a way that alters the 'hash' and changes the chain of evidence. Current chain of evidence procedures are in place and would therefore require updating rather than creation.



# Chain of custody operating model can also be enhanced to build trust in police digital sources

Policing will need to enhance chain of custody procedures in relation to evidential digital media to better assure veracity throughout an investigation to the point of presentation in prosecution proceedings, as well as to rebut any externally deepfaked media.



## Processes

*New governance processes and report will be required*

- Processes to manage and report new digital media assurance – i.e. ability to create a summary of all changes, such as cropping, to trusted digital media, and log of actions.



## Roles, skills & training

*Existing roles will need training on how trusted digital sources are assured and governed*

- An understanding of hashing technology and processes to provide greater assurance of digital media from the point of police ownership through to prosecution.
- This will include: Investigating officers who may be required to talk to such processes where these are examined in court settings; the CPS, judges and the prosecution / defense legal teams.



## Assurance, confidence & trust

*Digital fingerprinting will build confidence in trusted sources*

- Demonstrating and proving the nature of digital fingerprinting to the policing and justice system, as well as the public will increase assurance, confidence and trust in parts of the digital media evidence ecosystem. In turn this reduces the impact on the policing and justice system as trusted digital media sources can be used evidentially more effectively.
- Policing should also consider how deepfake technology could be used to alter / create media that resembles their own, such as BWV, and how such may intend to damage confidence and trust in policing, and / or implicate officers, which may have civil / criminal consequences (conduct matters).



## Ecosystem

*A wider ecosystem of trust will be developed*

- Working with industry may provide opportunities for tooling suggestions to be applied to media generation by non-police owned devices, such as Ring Doorbells. This could provide better veracity assurance for digital media gained via the public domain.

**The additional threat posed by deepfake technology being used in the tampering of digital evidence, means enhancing chain of custody procedures will be necessary to provide greater assurance that all digital evidence is unchanged from the point of police ownership, regardless of the media source and whether media may be a deepfake at the point of collection.**

# 04

## Testing digital media evidence in courts

A case study in deepfake-impacted procedures



# Deepfakes will cause disruption and new complexities as they are tested in courts

High profile cases and accusations involving deepfakes have been hitting the news for the past few years, from audio deepfakes of Scarlett Johansson, to well-established companies being defrauded for millions [1]. Cases with deepfakes at the heart of the alleged crime will inevitably increase as AI-enabled criminals become more sophisticated and prolific.

Additionally, deepfakes can be used to pervert the course of justice across all cases that involve use of digital media.

There are areas where existing protocols and processes will be impacted by use of deepfakes, and those where deepfakes will change the nature of a crime and therefore present new types of cases and application of new legislation.

*A Family case in the UK in 2019 saw a mother in a custody dispute use deepfake audio, after following instructions from online forums, to attempt to portray the father as aggressive. [2]*

## Deepfake-impacted procedures



The processes by which digital evidence is used in courts in different jurisdictions will need to change as new uncertainties emerge.

### This includes:

- CPS's evaluation of evidential value of digital media in making prosecution decisions
- Building cases using digital exhibits for defence/prosecution and claimants/ respondents
- How digital media is presented, argued and verified in courts

## Deepfake-enabled crime / threats such as



There are already new and evolving areas of legislation around the criminal use of deepfakes which will cause new types of case to come to courts.

### Such as:

- Use of AI tools for criminal activity
- AI generated illegal content including CSAM / intimate image abuse
- Use of AI generated images of others without consent

1. [The fake AI Scarlett Johansson is a reality check for Washington – POLITICO](#); Arup lost \$25mn in Hong Kong deepfake video conference scam (ft.com)
2. [Deepfakes in the courts | COUNSEL | The Magazine of the Bar of England and Wales \(counselmagazine.co.uk\)](#)

# Legislation must move fast to keep up with criminal use of deepfakes

There are laws currently in place to address criminal activities that are deepfake-enabled or deepfake-impacted.

FEB 1995

The creation, accession, possession, and distribution of CSAM (Child Sexual Abuse Material) is illegal, whether real or faked, as per the notation regarding pseudo-photographs ("*image whether made by computer-graphics or otherwise howsoever, which appears to be a photograph*") within the Protection of Children Act 1978, amended via s84 of the **Criminal Justice and Public Order Act 1994**.

The **Fraud Act 2006**, was amended to include 'Fraud by False Representation'. This amendment can apply to the use of deepfakes to mislead, or present untruths, for a perpetrators gain or induced loss on another.

JAN 2007

OCT 2023

'Perverting the course of justice' is an offence under Common Law in England and Wales. It can be applied to the ways in which deepfakes could impact the justice system, especially as this includes fabricating evidence. Common Law is "case law" applied by reference to previous cases / based on precedent. However, guidelines for judges and magistrates to assist in this sentencing came into place in October 2023 - this only applies to adults, and precedents may need to be set regarding deepfakes.

JAN 2024

The offence of sharing intimate images (including deepfakes) came into force (s188 of the Online Safety Act (OSA)) as an amendment introducing s66b Sexual Offences Act 2003.

The OSA was also amended to include a 'False Communications Offence' (s179) regarding the sending of messages known to be false information.

The creation of sexually explicit deepfakes is planned to be a new offence by UK government through an amendment to the Criminal Justice Bill\*. Under this new legislation, the **creation or design of intimate images** of another person that uses 'computer graphics or any other digital technology' to cause distress or harm will be considered a criminal offence. Not in force at Royal Assent: s15 (Foreign Interference) of the National Security Act 2023, in relation to criminalising foreign state backed misrepresentations intended to interfere with UK politics.

UPCOMING

*\*Due to the prorogation of parliament, the Criminal Justice Bill will not progress further at this time June 2024. Both Conservatives and Labour parties have pledged to criminalise intimate image abuse.*

However, deepfake specific legislation is limited; a victim of a deepfake may need to consider one, or a combination of existing laws, that are designed to protect other legal rights, to achieve justice. Continuing technology innovation is also likely to require the law to adapt.



# Challenging the veracity of digital media risks undermining procedures and overwhelming courts



01

## Increased scrutiny of digital media exhibits will add to complexity, effort and even harm

It is likely that use of digital media exhibits will be more frequently questioned as deepfakes by both defence and prosecutions as part of their case strategy.

For example: In cases of Serious Sexual Offence where the victim / perpetrator relationship is under scrutiny, deepfakes could exacerbate present issues such as missing messages on phones through casting further doubt on digital media exhibits.

The potential for additional examination of the authenticity of digital media may extend the time for digital forensic analysis of victims' devices, and delay court proceedings, causing further distress.

In addition, audio, video and images may be dismissed giving cases fewer evidential opportunities leading to potentially fewer convictions.

**Where the veracity of digital evidence is questioned, the complexity and length of proceedings and potential impact on victims may increase.**



02

## Volume increases driven by desire to demonstrate weight of evidence may overwhelm courts

The volume of digital evidence has already increased dramatically. As digital media becomes more challenging to use evidentially, this creates friction in the system with other corroborating evidence or higher volumes of digital evidence required to build cases, and further pressure created on disclosure obligations.

As prosecutors have already seen in the case of Body Worn Video (BWV), new ways to use digital media as evidence can significantly increase the volumes and effort involved in building and progressing cases through the court.

For deepfakes, it is conceivable that cases would bring a weight of digital evidence (for example a hundreds of similar photos or videos) rather than a smaller number to mitigate the risk of them being questioned as deepfakes.

**This would increase preparation, time and effort involved in case building and progression through the courts.**

## Different levels of rigour are likely to be applied to evidence in Civil and Criminal courts

While deepfakes will undoubtedly cause disruption to the justice system, it may impact and need to be addressed differently in the Civil and Criminal Courts.

### 03

#### Justice may be more readily undermined in Civil Courts

In many cases in the Civil Courts, there is not the same level of scrutiny of the evidence and therefore there may be greater risk of disruption and miscarriage of justice than the Criminal Courts.



I am increasingly worried about the evidence that will be put in front of Civil Courts as the entry level for evidence does not undergo the same level of scrutiny as Criminal Courts, and therefore, Civil Courts may be more susceptible to duping by deepfake evidence.”

**District Judge**

### 04

#### Criminal Courts will need to consider what constitutes ‘reasonable doubt’ in relation to digital evidence and how to deal with potential deepfakes

**Detection without understanding the source, tool or creation technique specifically will invite questioning of the reliability of the interpretation.**

For example:

- Would an image determined to be 85% likely of being a deepfake be deemed authoritative? What about 65%
- Would providing reasons behind this value provide more confidence? Would this additional information be easily understood by a jury?
- Would a competing tool with a lower confidence value negate the evidence?
- Would disputes over reliability of digital media – and different technical approaches to verification – end up undermining all digital media?

**This will create a period of disruption as court processes are tested and new precedents for case law are set.**

# Confidence in digital evidence chains will be tested



## Tooling

Deepfake detection tooling works in a variety of different ways, with equally diverse outputs. Common outputs are a statistical measure of likelihood of the media being a deepfake. Some tools provide further reasons behind this measure, others are less transparent. Further complications are introduced if multiple tools are used and provide differing outputs, or their outputs are combined.



## Interpretation & explainability

Expert interpretation – with the potential for disputes between experts - will become part of the process of reasoning whether digital media is a ‘deepfake’. Therefore, where digital media can be assured at point of production by trusted hardware, it is vital that such capabilities are put in place to reduce the burden on the justice system.



## Assurance, confidence & trust

The ability of a tool to detect deepfakes will need to be demonstrated. To do this effectively, all tools will need to be measured against common criteria. One important aspect of this will need to be assessing tools against a standard set of ‘test data’, where the number, type and quality of deepfakes in the test data is known. This will create a common benchmark for all deepfake detection tools to be used in the justice system. Without this it will be impossible to provide any level of confidence in the tooling.



# New expertise will be needed to challenge and defend digital media under scrutiny

## Policing digital forensics experts training in deepfake technology may act as expert witnesses



### DIGITAL FORENSICS EXPERT

Michael is part of the Digital Forensics Unit in West Midlands Police and specialises in identifying deepfakes and AI-generated content. He has a masters in Cybersecurity and holds several certifications in forensic analysis and AI technology. Michael works on high-profile cases, collecting and analysing critical evidence for investigations into serious and complex crimes. He is digitally savvy, an analytical thinker and has an excellent eye for detail. Due to his technical background, he also had a strong understanding of the analytical tools suite and when to apply each of them.

Michael plays a critical role in supporting investigations and then working with the prosecution team during the preparation of the case for court. Leveraging specialised tools and advanced methodologies, he analyses evidence packages to detect AI-generated images, videos and audio. His expert evaluations help determine the authenticity of evidence, ensuring only legitimate evidence is presented in court.

*Find out more about Michael and other future roles in the Appendices*



### Roles, skills & training

New roles, such as police digital forensics experts training in deepfake technology, will be needed to act as expert witnesses. They will need to be supported by skills and training.



### Ecosystem

Digital forensics experts will also be required from industry:

- From Tech / Social platforms – to provide expertise and insight into digital media on their platforms
- From subject matter experts – to provide insight into digital media authenticity, tools and deepfakes
- To offer defence teams access to similar expertise.



# 05

---

## Deepfake CSAM: a worrying new trend

A case study on deepfake-enabled crime



# Deepfakes present a new and challenging shift in the creation and proliferation of CSAM



Childlight recently published their inaugural 'Into the Light' Index [1] that aims to quantify the scale of Child Sexual Exploitation and Abuse (CSEA) globally for the first time. The figures are stark and disturbing:

**300M+**

children have been affected worldwide by online child sexual exploitation and abuse in the last 12 months

**7%**

of men in the UK report that they have engaged in online behaviours at some point in their lives that could be classed as online child sexual abuse offending

**1 in 8**

children globally have experienced non-consensual taking, sharing and / or exposure to sexual videos and images in the last 12 months

Childlight believe that CSEA should be treated with the urgency and priority of a global health emergency, such as a global pandemic.



Since early 2023, cases of offenders using generative AI to create child sexual abuse material (CSAM) has been increasing [2] and as tooling is easier to access and more realistic, it presents new avenues for offenders and new types of offending behaviours that ultimately present more avenues to harm to children.

The WeProtect Global Alliance's 2023 Global Threat Assessment set out three key challenges that the rise in deepfake CSAM posed:

- Difficulties distinguishing between deepfake and real-life content make it difficult to categorise reported imagery
- Irrespective of whether CSAM features "real children" or not, police must investigate each report to ensure a child is not being abused. As police currently lack the capabilities to automatically identify and triage for deepfakes, backlogs will grow and safeguarding will be delayed, prolonging children's suffering
- The consumption of deepfaked CSAM contributes to the 'market' this material and could change offender behaviours through perceptions of less harm, fuelling more extreme fantasies and lead to more active forms of abuse. It could also encourage a culture of tolerance for the increased sexualisation of children in the long-term.



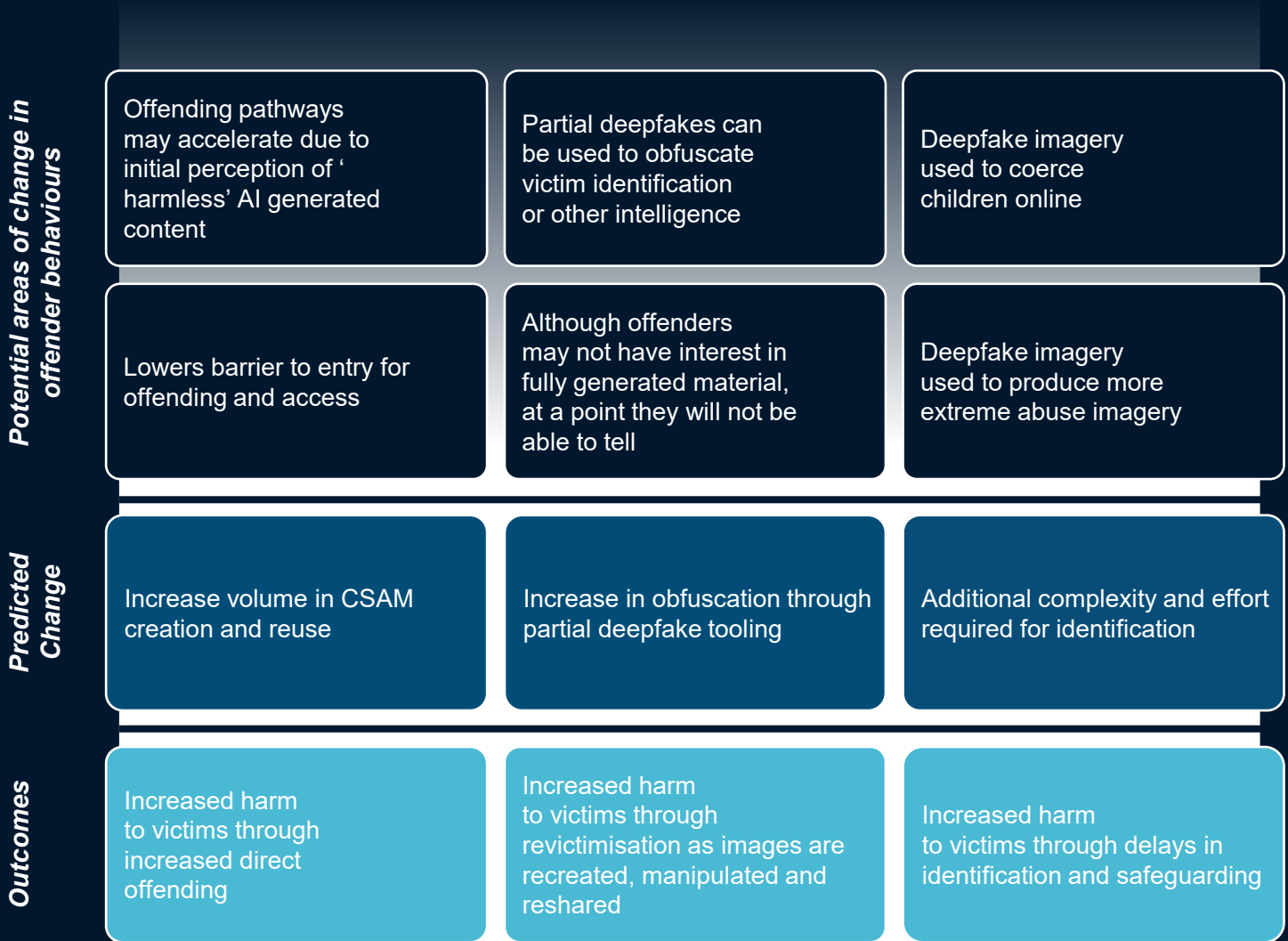
In October 2023, the Internet Watch Foundation published a report warning of AI-generated child sexual abuse images as a 'new and growing area of concern'. In one dark-web CSAM (Child sexual abuse material) forum alone over a one-month period, 20,000 AI generated images were found to have been posted, and around 50% of those were assessed as likely criminal images, taking 12 dedicated analysts a total of 87.5 hours. Around 3,000 images were finally judged to be criminal. [1]

**Any rise in this horrific crime; any additional complexity in safeguarding victims; any delays identifying offenders adds to an already unimaginable scale of harm to children.**

**We must understand and tackle the risk of deepfake CSAM as a priority.**

1. [Into the Light Index Reports | Childlight](#)
2. [Global-Threat-Assessment-2023-English.pdf \(weprotect.org\)](#)
3. <https://www.iwf.org.uk/about-us/why-we-exist/our-research/how-ai-is-being-abused-to-create-child-sexual-abuse-imagery/>

# Deepfake CSAM ultimately leads to increased harm and victimisation



# Offender techniques and use of detection tooling makes the response more challenging

Combinations of composite 'deepfake / real' images mean that detection may be challenging and making decisions more complex

**Real person / real background**



Current mode of CSAM images / video – investigators use tradecraft and cues to help them to identify victims and perpetrators

**Fake person / real background**



Deepfakes may be used to obfuscate a victim's identity and / or to victimise another child through adding their likeness to an abusive image.

**Real person / fake background**



Deepfakes can be used to obfuscate a background either fully synthetically or a different background leaving fewer opportunities for identification of location and victim.

**Fully deepfake – AI generated**



Fully synthetic images / videos skew people's view of what is harmful, lower barriers to offending – and may be training on CSAM

Rather than being confined to distinct categories, deepfakes exist on a continuum, ranging from subtle alterations to complete fabrications. This spectrum includes composites that merge authentic images to create believable scenes, manipulations that graft a real person's face onto a different body, and entirely synthetic creations that are generated from scratch.

The fluid nature of deepfakes means they can be tailored to varying degrees of realism, making the distinction between types more of a gradation than a clear-cut separation.

This variety of deepfakes necessitates a range of detection tooling, and presents a challenge to analysts who must 'grade' images according to severity and make decisions on criminality, for example of an image that is a child's face on an older looking body or vice versa.

## Detection of CSAM deepfakes introduce ethical and legal implications

The primary concern is the protection of children's rights and identities. Traditional detection methods can not be employed as freely, given the sensitive nature of the content, which limits the exposure of analysts to such material. This constraint necessitates the development of detection systems that are both highly effective and respectful of privacy.

## Clandestine distribution is a challenge to training detection systems

The illicit nature of CSAM deepfakes means that they are often distributed through clandestine channels, which makes them less accessible to researchers and law enforcement agencies working on detection algorithms. The secretive distribution networks also mean that there is a smaller dataset available to train detection models, making it harder to teach these systems how to identify CSAM deepfakes accurately.

## Offenders are motivated to continuously adapt their techniques

Additionally, the creators of CSAM deepfakes are often highly motivated to evade detection due to the severe legal consequences they face, leading to a continuous evolution of techniques to make these deepfakes more convincing and harder to detect. This creates an ongoing arms race, with detection methods constantly trying to catch up to the ever-improving creation techniques

Therefore, the detection of CSAM deepfakes is fraught with challenges that go beyond technical difficulties, encompassing ethical considerations, limited access to data, and the relentless innovation by malicious actors. It is a battle that requires not just technological sophistication but also a concerted effort from society to uphold the safety and dignity of the most vulnerable.



# A whole ecosystem response is needed to address this challenge

Organisations and governments, both nationally and internationally, tackling this challenge are beginning to see how deepfake capabilities are impacting offending behaviours and risk to children, how to engage with the complexity of legislating this area and identifying the gaps in the existing shared solutions.

## Ecosystem



- Agree a shared taxonomy and language to accurately discuss and share information about deepfakes in this area. This is a critical priority for enabling other collaborative work in this area

## Roles, skills & training



- New training to develop to interpret and understand deepfake tooling
- Establish a knowledge base / community that includes detection historical tooling
- Best practice in detection and tooling should be shared to support development of skills and training

## Policy & Regulation



- Build clarity and consensus on legislative ability
- The legislative landscape must be scrutinised to understand its ability to respond to new complexities

## Processes



- Investigate automated triage processes now to manage future volumes

## Tooling



- Tools / tool suite to speed up identification and support subjective decision making
- Design an operating model for adaptation and changes to source tools and detection tooling

## Interpretation & explainability



- Agreed new metadata fields for tagging
- New hashing metadata fields need to be agreed between organisations to assist analysts / officers in understanding i.e. the extent of identified deepfake image, the source tooling and analysis tooling used for decision making

## Human behaviours & understanding

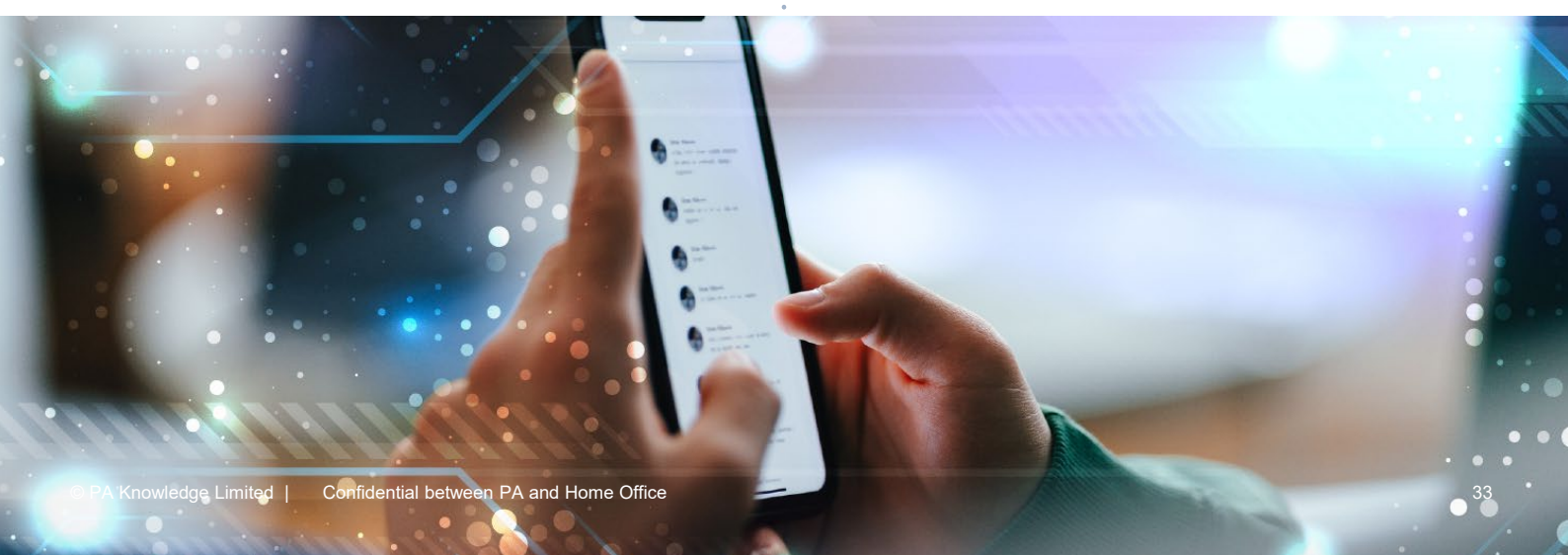


- Research bias and behaviours
- In decision making as well as emerging offending behaviours need to be further understood

## Assurance, confidence & trust



- Share explainability between organisations
- Agreement on any required changes to classification of material



# We must act now to build the right foundations to tackle the evolving threat

**Initial steps must include building collective agreement and adaptation of existing processes quickly to address the emerging challenges. Without this foundational clarity in place, it will be impossible to drive an effective and fast collective response.**

**Priority recommendations:**

**01**

Agreement on prioritisation of challenges and definition of collective taxonomy

**02**

Agreement on 'quick wins' across government, law enforcement and industry – for example:

- New hashing metadata fields to be agreed
- Agreement on classification models
- Testing opportunities and effectiveness of existing tooling options (i.e. can AI tooling be used to identify deepfake 'sets' more efficiently?)

**03**

**Understanding of current and future requirements for capability across the ecosystem, including:**

- System to be stress tested (i.e. wargaming / scenario planning) against sudden increases in volume to mitigate increased harm to victims
- Exploration of management of increasing volumes, use of sets and other specific system characteristics

**04**

Best practice in detection and tooling should be shared where they are already being used and tested

**05**

Initial awareness and skills development put in place to enable different roles to understand and respond to how the change in threat impacts their work – for example training in decision making where deepfakes can cause 'grey areas' in legality (i.e. mixture of apparent 'ages' in composite image).

**06**

Working with third sector and other organisations to monitor changes to offending behaviours and engagement with CSAM deepfakes

# 06

## Online Fraud: Empowering the public through education

A case study on deepfake-enabled crimes



# Cybercrimes will continue to grow exponentially as they become enabled by deepfakes

Email phishing scams are the most prevalent scams on the internet in which criminals masquerade as a known or trustworthy individual to gain the trust of a victim. Audio, image and video deepfakes are already considered a fast-growing threat ('deepfake fraud') that gives cyber criminals more manipulative and sophisticated tools to be able to dupe their victims.

## Phishing, Spearphishing and Spoofing: from email scam to deepfake phishing



Peter is a retired CFO. He is on the board of two charities for children with special educational needs, a cause close to his heart. He thinks of himself as a savvy, smart man – if a little out of touch with the kids.

One day he receives a panicked voice note from his granddaughter who is on her gap year travelling in Thailand. She's been in a moped accident, she's not badly injured but needs £2000 to pay for the hospital or she can't get treated.

Peter quickly sends the money via the link that she's whatsapp'd over.

**Peter has been scammed - Audio is one of the easiest areas to convincingly deepfake**

## Online dating scams / romance fraud



Jackie is a single parent who works as a hospital administrator. She has a busy life balancing work, checking in on her elderly mother and looking after her 4 year old boy.

Jackie has grown close to David, a man she met online who lives in Greece. They message frequently, share selfies and photos and video call every few weeks. Over the past year, Jackie has sent David around £10,000 to help him with his business and emergency costs.

**Jackie is being scammed – scammers may use a convincing mix of video and image deepfakes**

## Next generation 'business email compromise'



Piyush, a finance executive for a large engineering firm gets an invitation for a video call with his international colleagues. Although initially suspicious, once he joins the call he sees the CFO and a number of other colleagues online. After an in-depth discussion about international cash flow, he transfers £20 million in 15 separate transactions.

This scenario is based on the real-life scam that defrauded the design and engineering firm Arup of millions. [1]

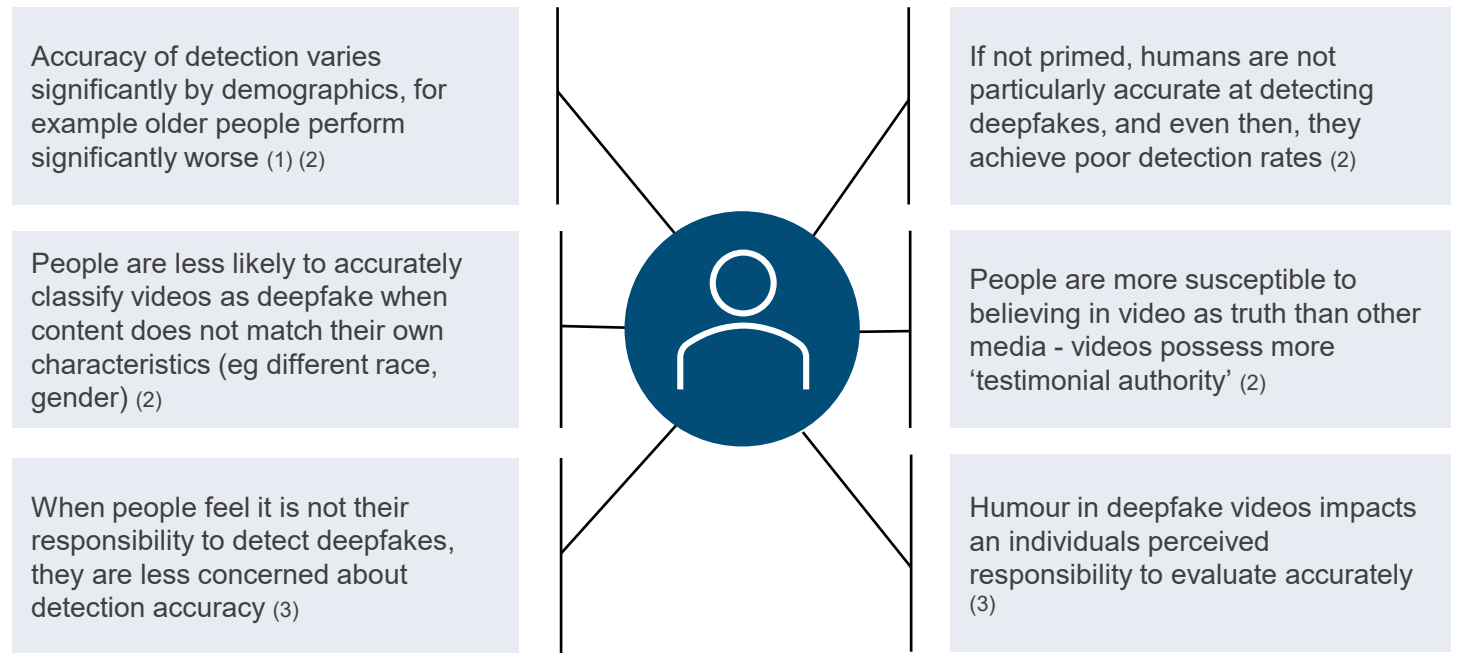
**Piyush was scammed – this sophisticated attack used social engineering, deepfake live-streams and audio**

1. <https://www.ft.com/content/b977e8d4-664c-4ae4-8a8e-eb93bdf785ea>




# To equip people with preventative tooling, we need to understand human behaviour

**Effectively perpetrating cybercrime relies on gaining the trust and confidence of an individual, and it is made exponentially more effective through deepfakes.**

Studies reveal that people are very poor at detecting deepfakes and have deep-seated biases that impact accurate detection.



As with other modes of cyber crime, such as text scams and bank scams, deepfakes as a risk should be integrated into broader formal and public education campaigns:

- Early integration of the topic into education

- Public education campaign and tools to suit different demographics

- Working with industry to add deepfakes to existing message – i.e. in banking apps


1. Mumford, Jeremy, Brigham Young University (2024), *Improving Human Recognition of Deepfakes (Undergraduate Honors Theses)*
2. Lovato, J., St-Onge, J., Harp, R. et al. (2024) *Diverse Misinformation: Impacts of Human Biases on Detection of Deepfakes on Networks*
3. Stuart Napshin, S., Jomon Paul, Justin Cochran, (2024) *Cyberpsychology, Behavior, and Social Networking*

# Our 'Deepfeed app' future concept helps to safeguard the public from advanced deepfake exploitation

Imagine an app that is adaptable to different people's needs, co-developed by government and industry, supported by tech platforms and can help people make better critical decisions to detect deepfakes.

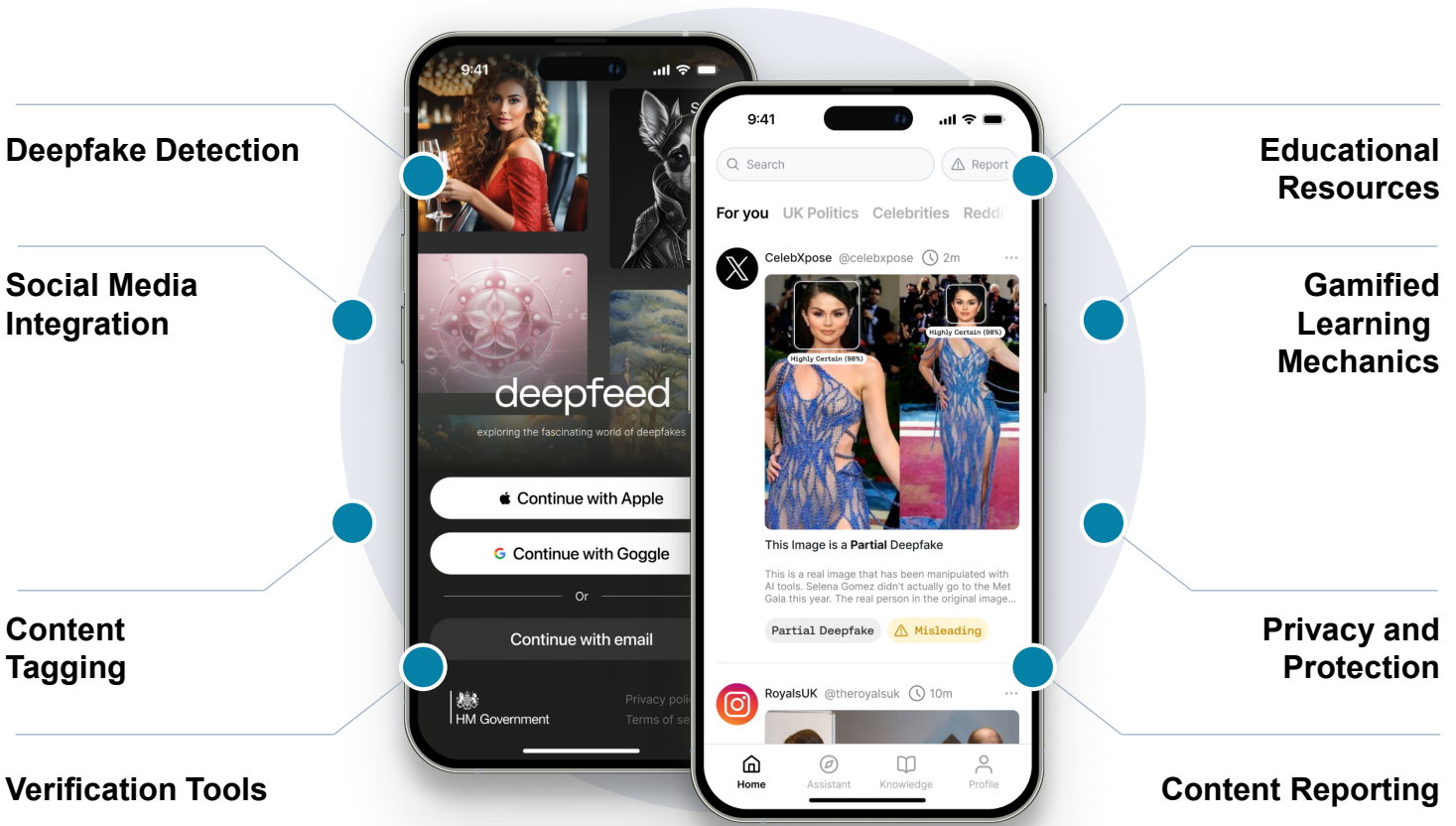
Introducing Deepfeed, a cutting-edge application dedicated to exploring the fascinating world of deepfakes in a trusted environment. Deepfeed aims to demystify deepfake media, helping users identify and understand the deepfakes they encounter online, preventing them from become victims of scams using deepfake content.

Deepfeed aims to support the public in the following ways:

**01 Education and Awareness**  
Equipping users with knowledge about the risks of deepfakes and provide practical tools and tutorials to help them identify and respond to potential threats effectively, fostering a vigilant and informed community.

**02 Detection and Verification**  
Utilising tooling to empower users in distinguishing between authentic and manipulated media, enhancing their ability to navigate and mitigate the impact of sophisticated deepfake content.

**03 Community and Support:**  
Creating an environment where users can discover deepfake-related content, learn through experiences, and report suspicious content, creating a proactive community that collectively safeguards against the spread of misleading/harmful material.



Our Deepfeed concept is an innovation tool designed to stimulate discussion and thinking

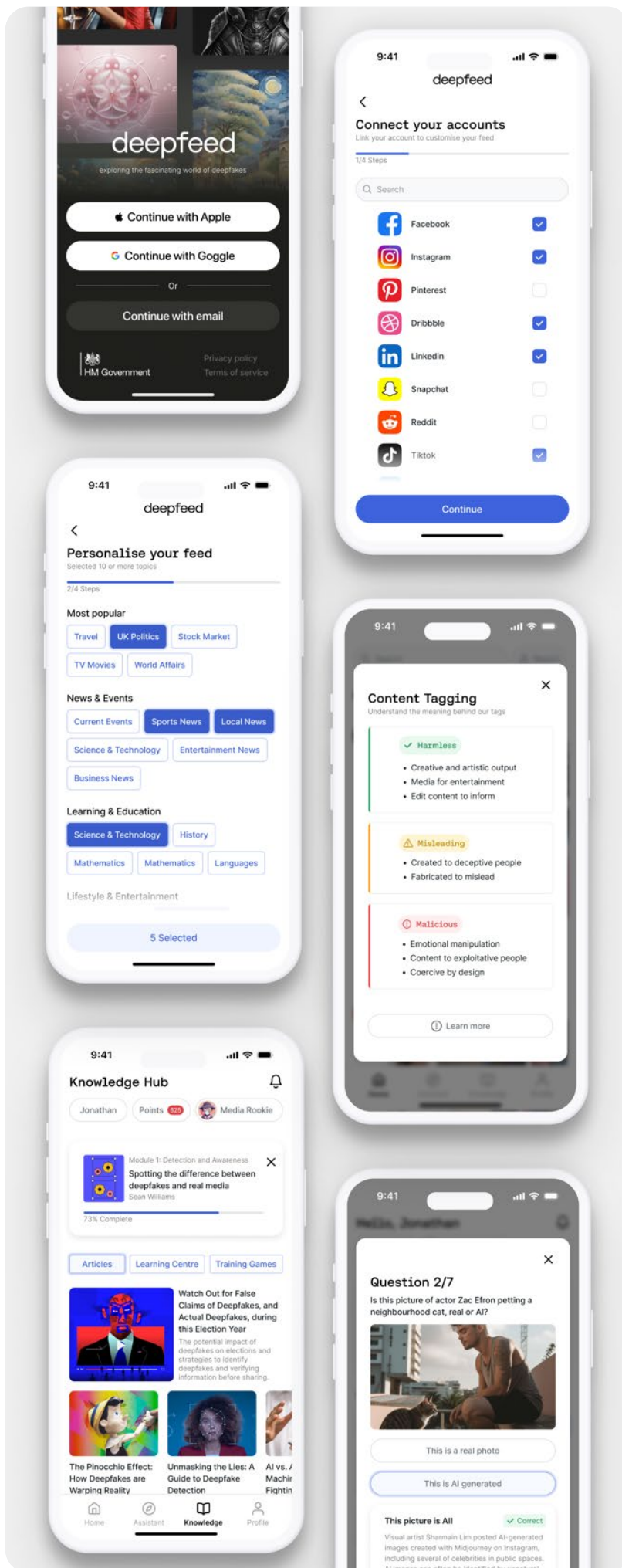
# 01 Education and Awareness

**Deepfeed educates and creates awareness by integrating with users' social media accounts, creating a real-time insight into their content and feeds.**

Deepfeed serves as a knowledge hub offering comprehensive articles, learning modules, and interactive training games designed to empower users with the skills and awareness needed to detect and respond to deepfake threats effectively.

- **Social Media Integration:** Integrating with users' social media accounts, generating a personalised, real-time content feed that can help to identify deepfakes within their networks.
- **Detailed Content Tagging:** Employing the latest tooling and sharing across the platform ecosystem to provide content tagging, systematically categorising, and flagging deepfake-related content across various media types. Users can identify and understand the nuances of deepfake manipulation, empowering them to make informed decisions about the media they encounter.
- **Gamified Educational Resources:** Enhancing user engagement through gamified educational resources, such as interactive modules, articles and training games. Designed to simulate real-world scenarios involving deepfakes, allowing users to practice detection skills in a safe environment while reinforcing their understanding of the technology's implications. This approach not only educates users effectively but also encourages active participation and continuous learning.

**Our Deepfeed concept is an innovation tool designed to stimulate discussion and thinking**



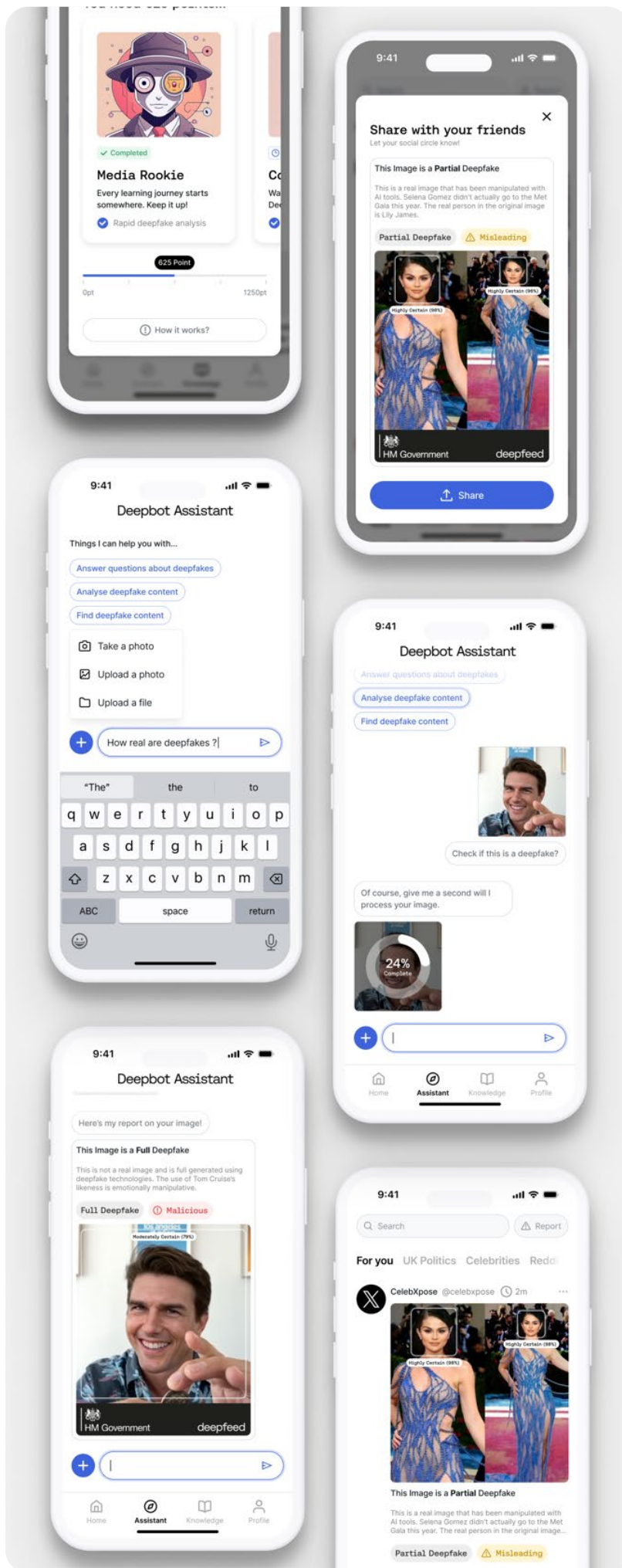
## 02 Detection and Verification

**Deepfeed detects deepfake media through scanning, cataloging and tagging social media content, providing users with shareable detailed explanations on potential deepfakes.**

It also features an 'AI assistant' that guides users through questioning the authenticity of content, offering options to upload or share media links for analysis and verification, ensuring users can confidently discern between genuine and manipulated content.

- **Deepfake Detection Algorithms:** Utilising advanced deepfake detection algorithms that scan, catalog, and tag social media content. These algorithms analyse visual and audio cues, identifying anomalies and patterns indicative of deepfake manipulation. This process ensures that deepfake content is promptly flagged to a user.
- **Verification Tools:** Providing users with a chat assistant interface which guides them through authenticating media content, enabling direct uploads or links for analysis with advanced algorithms. Interacting with the chat assistant clarifies content authenticity, empowering users to confidently distinguish between genuine and manipulated media.
- **Shareable Verified Media:** Generating detailed explanations and evidence-backed assessments that users can share confidently. These shareable reports include explanations of detection methods used, highlighting specific manipulations or inconsistencies found. Educating users and raising awareness and discussion about deepfake threats within their online communities

**Our Deepfeed concept is an innovation tool designed to stimulate discussion and thinking**





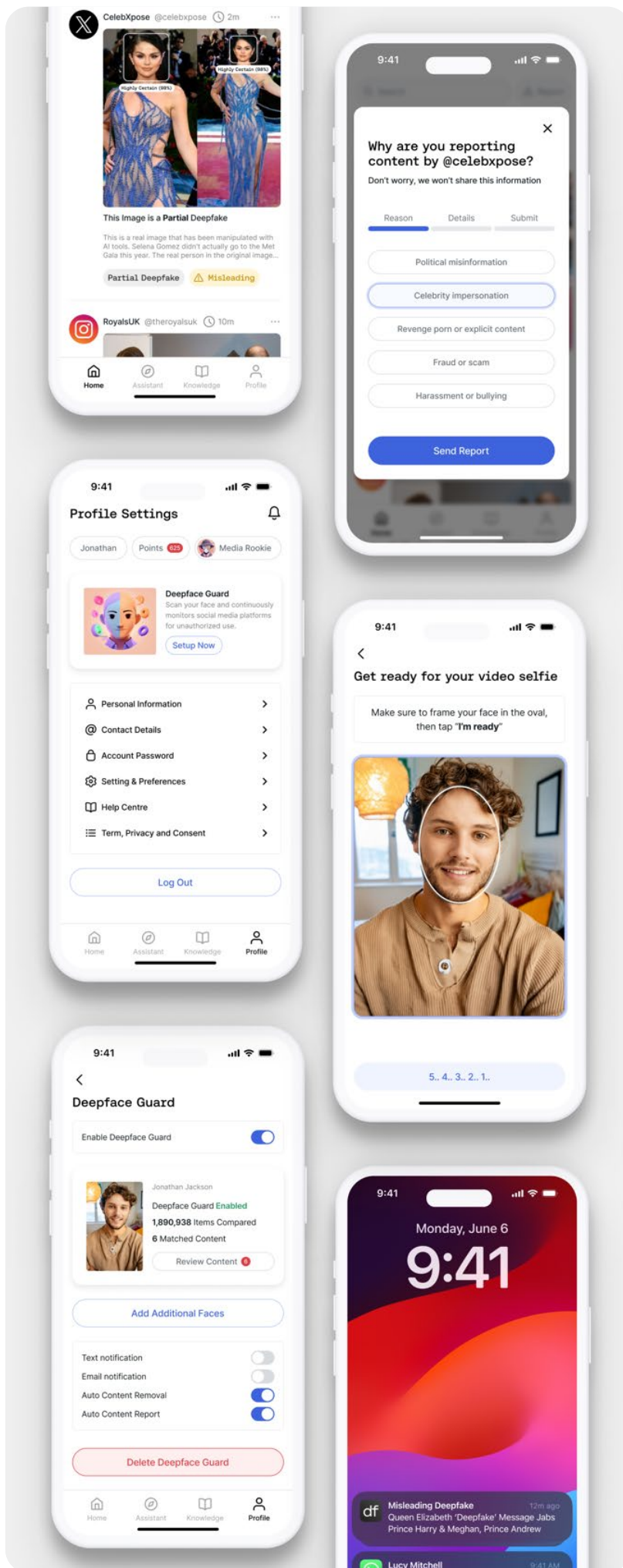
## 03 Community and Support

**Deepfeed creates community and support by offering robust automatic and manual content reporting mechanisms, ensuring swift action against illegal deepfake content or legal deepfake content that platforms may want to tag.**

It prioritises privacy and protection, offering tools to prevent unauthorised use of individuals' likenesses in deepfakes. Additionally, Deepfeed provides real-time alerts to notify users when harmful or misleading deepfake content is detected on their social media accounts and other online platforms, empowering them to stay vigilant.

- **Content Reporting:** Implementing both automatic and manual content reporting mechanisms, allowing users to flag deepfake content. This proactive approach facilitates community involvement in identifying and addressing potential misleading or illegal content, ensuring a safer online environment for all users.
- **Privacy and Protection:** Offering robust tools designed to safeguard users' privacy and prevent unauthorised use of their likeness in deepfakes. Users are notified upon detection and given options to report and remove identified content, empowering them to maintain control over their digital identities.
- **Real-time Alerts:** Delivering real-time alerts to notify users instantly when there is harmful or misleading deepfake content detected on their social media feeds. These alerts provide timely warnings and actionable insights, providing users with the correct information to protect themselves from misinformation and manipulation whilst online.

**Our Deepfeed concept is an innovation tool designed to stimulate discussion and thinking**



# Deepfeed is a concept, but it highlights the need and complexity in public education for prevention

**The online landscape of digital media, public and private platforms and human engagement is vast and complicated.**

Public education, as a method of prevention, allows targeted support for individuals as a type of intervention that can help people use available tooling to support critical thinking and come to better decisions to protect themselves.

## 01

Although it will not stop the creation and malicious use of deepfakes, education is key to helping individuals protect themselves – from early school education and throughout life.

## 02

**To educate effectively we need to understand human bias, behaviour and fallibilities, as well as how people interact with deepfakes now**

- Research and insight into how people respond to deepfakes and the effectiveness of recent education measures (such as Covid misinformation tagging)
- Create 'customer segments' for different cohorts of the public, particularly in relation to vulnerability

## 03

**A public education response needs to be coordinated across government**

- Create initial strategy, plan and guidance on deepfake education and be prepared for it to change quickly and iteratively
- Clarify ownership of deepfake education and innovation

## 04

The wider ecosystem will be critical to creating real impact

- Work with industry to identify models for public support and education that can be updated– i.e. banking warnings on money transfers.
- Working with 3rd sector organisation, platforms and companies to assess where tooling and information can be provided to the public
- Solutions can be co-developed with the private sector

## 05

Any education solution needs to be adaptable to different needs and cohorts

- Different segments of the population will require different approaches, i.e. from websites to different app 'skins' and features.

## 06

Start testing and learning now to keep up with a fast-changing landscape

- Test initial tooling and gamification to support deepfake education for key vulnerable cohorts – this can be detection and education tooling, such as gamification of deepfake detection.

# 07

## Conclusion



# Technology might not have all the answers, but that doesn't mean that answers can't be found

**Like the advent of the internet and the digital revolution, AI and all emerging technologies will impact how crimes are committed, how the law is enforced and how citizens are protected from harm.**

Technologists predict that within 12-18 months, humans and many machines **will not be able to detect deepfake media images, videos and audio from authentic digital media**, and that alone a single tool, or even a suite of tools will likely never provide a complete and definitive answer.

However, **there is still a critical role that tooling needs to play in assisting humans to detect deepfakes**, as doing so will support decision making and provide a vital way of adapting 'deepfake-impacted' procedures and combatting 'deepfake-enabled' crimes.

Tooling must be coupled with a wider ecosystem and operating model designed with this new paradigm in mind. Without this holistic approach, tooling will not be fit for purpose or effective in supporting human decision making, particularly where real-world consequences in law enforcement and government can be the most extreme.

In analysing a range of use-cases and considering the impacts of deepfakes, new operating models required to enable humans to effectively use tooling to detect deepfakes, we have been able to recognise common themes that must be considered when planning a response to this challenge:

## 01

### *Assurance*

Using deepfake tooling will require clarity, governance and assurance built-in to decision making processes which will either be based on a specific 'threshold' of outputs or that rely in human judgement. Governance will be needed to create standards around use of tools and assurance where AI algorithms are used.

## 02

### *Understanding the risk*

We will need to consider each threat / crime type, its specific context and processes to identify where effective interventions can be applied, as well as interrogate the risk appetite of each organisation in relation to the context of the threat, the specific decision and particular roles involved.

## 03

### *Designing for people*

We must understand human biases and behaviours in relation to deepfakes to adequately provide and use tooling effectively

## 04

### *Ongoing adaptability*

Long term and ongoing adaptability must be central to any solutions to keep pace with the existing deepfakes 'arms-race' and to prepare to adapt to future emerging technologies that will further disrupt the landscape, such as quantum computing.

# Innovation, ingenuity and collaboration are critical to putting adaptable solutions in place

To quickly and systematically adapt to emerging technologies, we will need to:

- 01 Establish a clear **shared taxonomy and framework** through which we can accurately describe, assess and tackle the challenge of deepfakes
- 02 **Prioritise key missions** where the impact and risk of increased harm is already becoming apparent (i.e. CSAM)
- 03 **Establish a roadmap for change** across policing capabilities to gain clarification and consensus on direction of travel, prioritisation and what is needed to tackle the challenge, including a framework for governance and assurance that can support innovation and testing of tooling solutions.
- 04 **Engage in a full ecosystem** approach, building on work such as the Deepfake Challenge to work with industry, platform providers, manufacturers and the third sector to collectively bring deepfakes up the agenda
- 05 **Look for quick win opportunities** such as:
  - Specific intervention opportunities to test and learn – for example in testing tooling types and combinations across pilot areas in policing and CSAM
  - Opportunities for building public awareness into existing campaigns – such as engaging with banks to include it in fraud information and awareness
  - Assess where interventions may be required now to existing programmes that are actively developing digital capabilities that will be impacted by deepfakes.



# A

---

## Appendix A - Personas



# We've imagined how key users might interact with tooling in the future to detect deepfakes

This set of Appendices contains:

**A:** Personas for policing roles that explore the needs and challenges of working with deepfakes in the near future



**Police Constable Olivia**

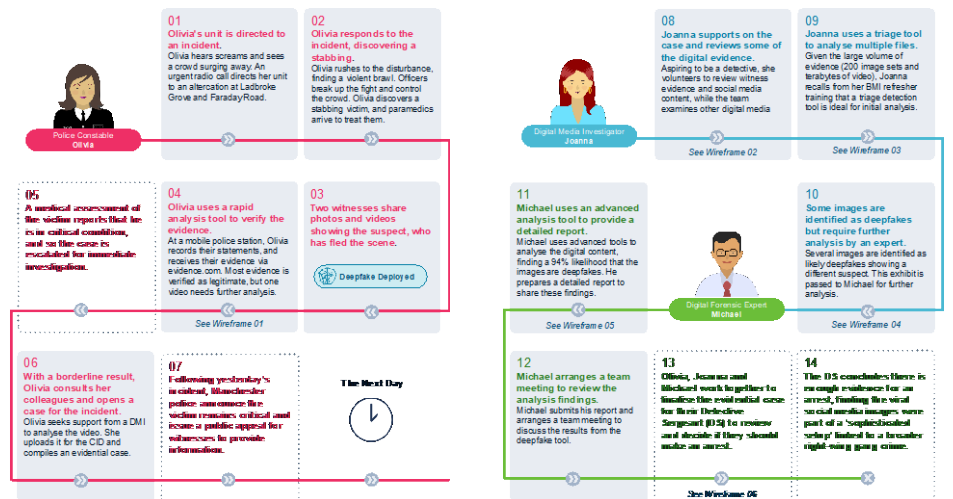


**Digital Media Investigator Joanna**



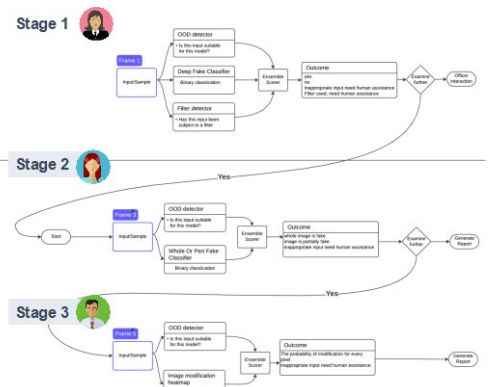
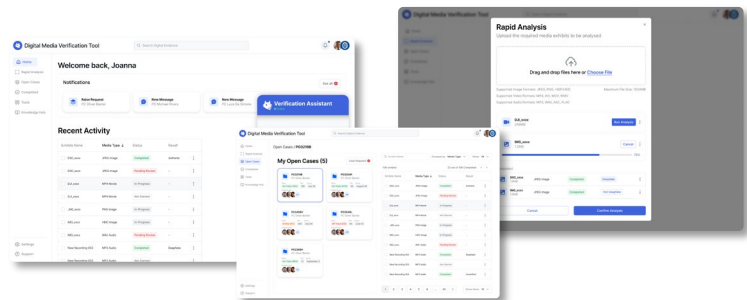
**Digital Forensics Expert Michael**

**B:** A scenario designed to explore new ways of working, interventions, assessments and processes in response to a fictional incident involving deepfakes at a live event



**C:** Exploratory wireframes to illustrate how tooling could function for different roles and demonstrating some of the challenges of tool usage and interpretation.

In addition, Advai have developed a model for a technical solution to enable the proposed deepfake tooling.



# Police Constable, Olivia

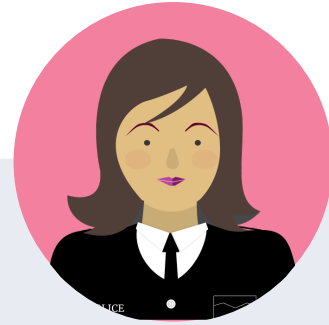
**Olivia, an experienced officer with West Midlands Police for 10 years, has strong legal knowledge and specialises in investigating crimes like harassment, disorderly conduct, and shoplifting. While aware of evidence analysis tools, she is actively training to improve her proficiency in new technologies.**

## Current Pain Points

- High workload: Handling multiple live cases simultaneously.
- Technical proficiency: Olivia requires comprehensive training to effectively utilise and understand digital technologies critical to modern law enforcement.
- Time pressure: Working under tight deadlines, especially under the PACE clock during arrests.
- Accuracy: Ensuring evidence legitimacy to prevent wrongful arrests or overlook threats to public safety.

## Key Needs

- With deepfake crime on the rise in the UK1, Olivia needs deepfake analysis tools that are quick, portable and user-friendly
- Training and upskilling to keep updated with the latest technologies and tools
- Access to digital/specialist support, or more advanced tools, where the likelihood of deepfake is ambiguous
- Understand the factors that are taken into consideration as to how likely digital content is real or fake
- Use software that is reliable and accurate so she can confidently present the information to the Sargent and other parties involved



## Role

Olivia is a frontline law enforcement officer maintaining public order.

## Responsibilities

- Secure crime scene, collect evidence (including exhibits), search items, interview suspects, and take witness statements.
- Use specialist deep fake analysis tools to assess evidence authenticity.
- Decide on suspect arrest and police custody, conducting searches and cataloging personal belongings.
- Note ambiguous evidence for further investigation.
- Construct evidential case, highlighting strengths and weaknesses for decisions such as caution, charges, or no further action (NFA).
- Compile detailed case file with all relevant information, evidence, and recommendations for review

## Tools Used

- 'Lite' Video Analysis Tools: Detect inconsistencies in video footage.
- 'Lite' Audio Analysis Tools: Identify manipulated audio.
- 'Lite' Image Analysis Tools: Examine image authenticity at the crime scene.



# Digital Media Investigator, Joanna

Joanna is a seasoned frontline officer who has undergone regular DMI training over the past 5 years, including advanced courses in deepfake analysis and attends bi-annual refresher courses to stay updated with the latest advancements in technology and cybercrime trends. Joanna's expertise allows her to efficiently handle digital evidence and advise on complex investigations.

## Current Pain Points

- Evolving threats: Joanna has to constantly adapt to new and sophisticated methods used by criminals involving deep fakes – constantly adapting to new tools & upskilling
- Complex evidence: Impact of dealing with deepfake content may require, more effort and resources
- Bias: fair interpretation, degree of subjectivity in decision making
- Decision making: Knowing when the case needs to be passed on to the Digital Forensic Expert

## Key Needs

- Quick and sophisticated deepfake detection tools needed.
- Training required on latest deepfake analytical tools and result interpretation.
- Attend bi-annual DMI refresher courses on deepfake trends.
- Stay updated on legislation impacting digital crime.
- Access to tools for generating detailed reports on inconclusive investigations handed to digital forensic experts.



## Role

Joanna assists investigations that need specialised digital investigative support.

## Responsibilities

- Conducts open and closed source deepfake investigations to support incidents and operations.
- Retrieves exhibits, including social media videos and audio recordings, via evidence.com.
- Utilises specialised tools for initial analysis of exhibits.
- Prepares investigation materials and communicates with Crown prosecutors.
- Shares best practices and provides training on deepfake content regionally and nationally.

## Tools Used

- Video Analysis Tools: Detect inconsistencies in video footage (e.g., Sentinel, Deepware).
- Audio Analysis Tools: Identify manipulated audio (e.g., Adobe VoCo, Descript).
- Image Analysis Tools: Examine image authenticity (e.g., Forensically).
- Quantitative Analysis Tools: Perform data analysis and reporting (e.g., EnCase).

# Digital forensics expert, Michael

**Michael, based in the Digital Forensics Unit of West Midlands Police, specialises in detecting deep fakes and AI-generated content. With a master's degree in Cybersecurity and multiple certifications in forensic analysis and AI technology, he contributes crucial evidence to high-profile cases involving serious and complex crimes destined for trial. Michael is digitally adept, analytically astute, and possesses a keen eye for detail, leveraging his technical expertise to skillfully apply analytical tools as needed.**

## Current Pain Points

- Adapts to evolving and sophisticated criminal methods in deepfake technology.
- Stays updated with latest tools and technologies through continuous upskilling.
- Manages large volumes of evidence, ensuring thorough analysis under tight deadlines.
- Maintains awareness of potential biases in evaluating deep fake likelihood to minimise subjectivity.

## Key Needs

- Receive correct evidence package via digital evidence portal.
- Easily navigate well-structured, tagged, and metadata-rich evidence for efficient workflow.
- Access and apply specialised tools for evidence analysis.
- Attend regular AI training to stay current with evolving technologies.
- Interpret and communicate qualitative and quantitative data effectively.



## Role

- Michael analyses evidence to detect AI-generated content, ensuring only legitimate evidence is presented in court.

## Responsibilities

- Receives evidence packages including social media videos, audio recordings, and doorbell footage.
- Utilises specialised tools to analyse exhibits, eliminating potential biases.
- Generates detailed reports with technical descriptions of anomalies detected.
- Provides expert evaluation on the authenticity of evidence.
- Submits final report to support case preparation for court proceedings.

## Tools Used

- Video Analysis Tools: Detect inconsistencies in video footage (e.g., Sentinel, Deepware).
- Audio Analysis Tools: Identify manipulated audio (e.g., Adobe VoCo, Descript).
- Image Analysis Tools: Examine image authenticity (e.g., Forensically).
- Quantitative Analysis Tools: Perform data analysis and reporting (e.g., EnCase).

# B

---

## Appendix B – User Scenarios



# Example scenario of a utilising a deepfake verification tool during day two of Notting Hill Carnival

Notting Hill Carnival is buzzing with parades and live music. Olivia, a Met Police Constable, is stationed at Ladbroke Grove. Briefed on heightened security due to gang tensions and potential disruptions, she monitors the crowd for trouble along with other officers.



Police Constable  
Olivia

## 01 Olivia's unit is directed to an incident.

Olivia hears screams and sees a crowd surging away. An urgent radio call directs her unit to an altercation at Ladbroke Grove and Faraday Road.

## 02 Olivia responds to the incident, discovering a stabbing.

Olivia rushes to the disturbance, finding a violent brawl. Officers break up the fight and control the crowd. Olivia discovers a stabbing victim, and paramedics arrive to treat them.

## 05 A medical assessment of the victim reports that he is in critical condition, and so the case is escalated for immediate investigation.

## 04 Olivia uses a rapid analysis tool to verify the evidence.

At a mobile police station, Olivia records their statements, and receives their evidence via evidence.com. Most evidence is verified as legitimate, but one video needs further analysis.

## 03 Two witnesses share photos and videos showing the suspect, who has fled the scene.



Deepfake Deployed

See Wireframe 01

## 06 With a borderline result, Olivia consults her colleagues and opens a case for the incident.

Olivia seeks support from a DMI to analyse the video. She uploads it for the CID and compiles an evidential case.

## 07 Following yesterday's incident, Manchester police announce the victim remains critical and issue a public appeal for witnesses to provide information.

The Next Day



## Example scenario continued

Joanna, a frontline officer with five years of training in digital media investigation, is assigned a new case. As an ambitious detective, Joanna eagerly volunteers to examine the witness evidence and social media content gathered from the second day of the Notting Hill Carnival.



Digital Media Investigator  
Joanna

### 08 Joanna supports on the case and reviews some of the digital evidence.

Aspiring to be a detective, she volunteers to review witness evidence and social media content, while the team examines other digital media



See Wireframe 02

### 09 Joanna uses a triage tool to analyse multiple files.

Given the large volume of evidence (200 image sets and terabytes of video), Joanna recalls from her BMI refresher training that a triage detection tool is ideal for initial analysis.



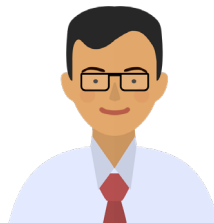
See Wireframe 03

### 11 Michael uses an advanced analysis tool to provide a detailed report.

Michael uses advanced tools to analyse the digital content, finding a 94% likelihood that the images are deepfakes. He prepares a detailed report to share these findings.



See Wireframe 05



Digital Forensic Expert  
Michael

### 10 Some images are identified as deepfakes but require further analysis by an expert.

Several images are identified as likely deepfakes showing a different suspect. This exhibit is passed to Michael for further analysis.



See Wireframe 04

### 12 Michael arranges a team meeting to review the analysis findings.

Michael submits his report and arranges a team meeting to discuss the results from the deepfake tool.



### 13 Olivia, Joanna and Michael work together to finalise the evidential case for their Detective Sergeant (DS) to review and decide if they should make an arrest.



See Wireframe 06

### 14 The DS concludes there is enough evidence for an arrest, finding the viral social media images were part of a 'sophisticated setup' linked to a broader right-wing gang crime.



# C

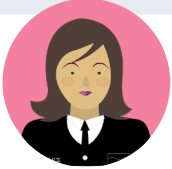
## Appendix C – Wire Frames & technical



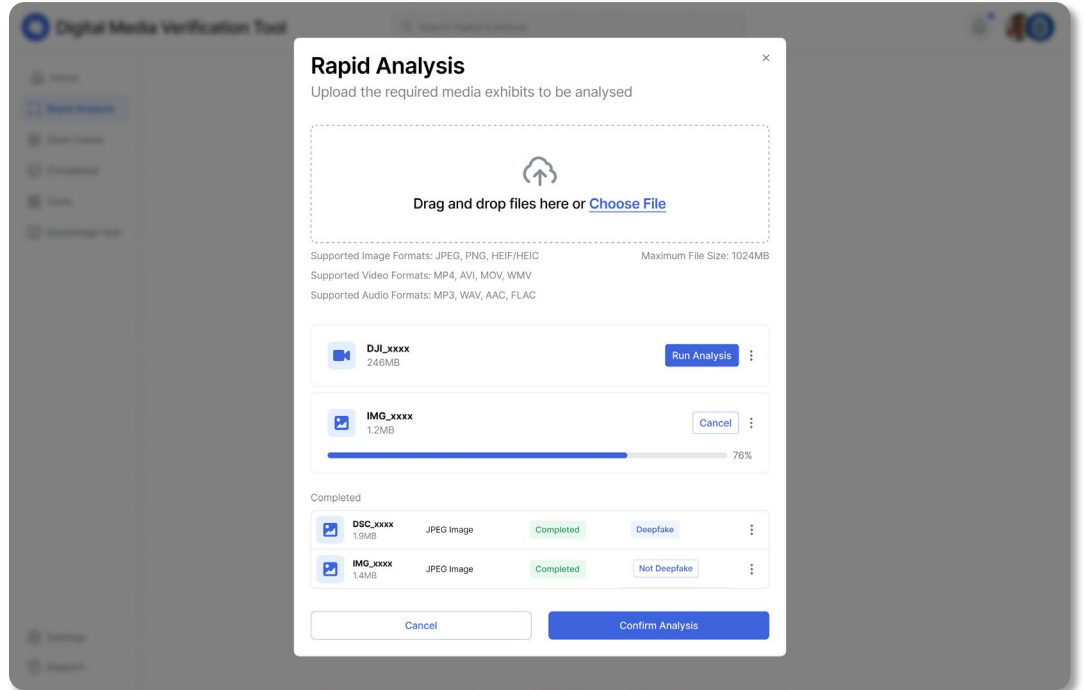
# Digital Media Verification Tool: Conceptual high-fidelity wireframes for evidential police investigations

## Wireframe 01 Rapid Analysis Tool for Deepfake Detection.

A tool designed to provide a swift preliminary assessment of digital media to determine its authenticity. It will quickly analyse the media and give the user an initial read out of whether the content is likely a deepfake or if it requires further, more detailed analysis.



- Used on the frontline to assist initial evidential assessment
- Uses preselected tools and techniques to give a risk indicator

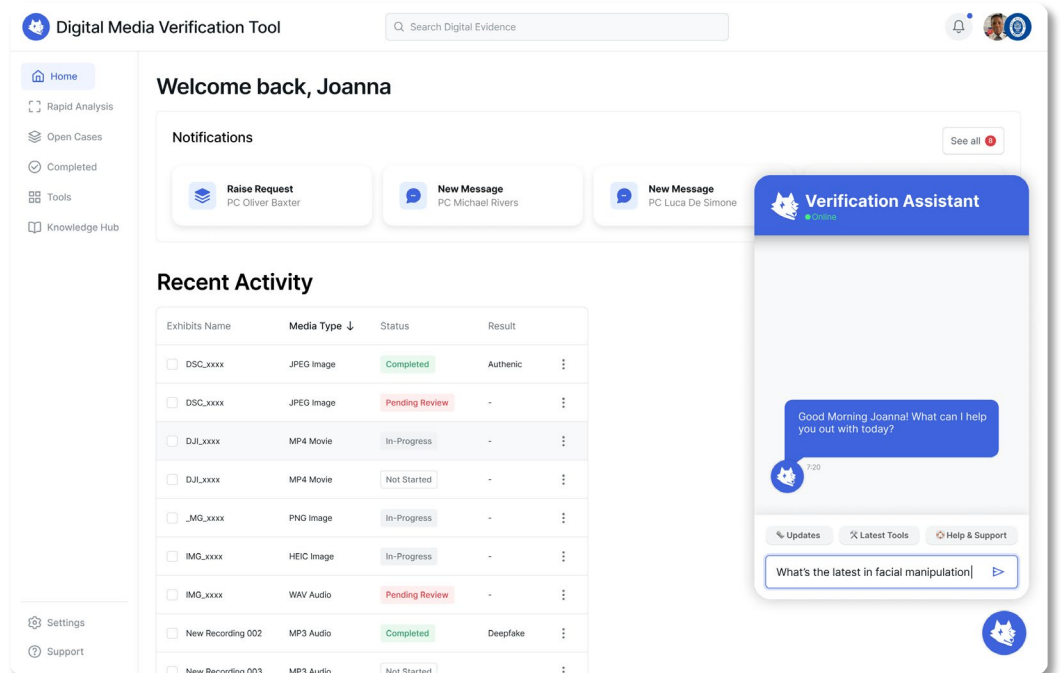


## Wireframe 02 Digital Forensic Dashboard

The dashboard is a central hub for law enforcement to verify digital media evidence. It features advanced tools for deepfake detection, metadata examination, and reporting, all in a unified interface. An integrated AI assistant chatbot supports officers by providing guidance, answering queries, and offering real-time insights to enhance investigative efficiency and accuracy.



- Used for more advanced analysis by DMIs to make digital media assessments
- AI assistant would help the DMI stay up to date and select the right tooling and approaches



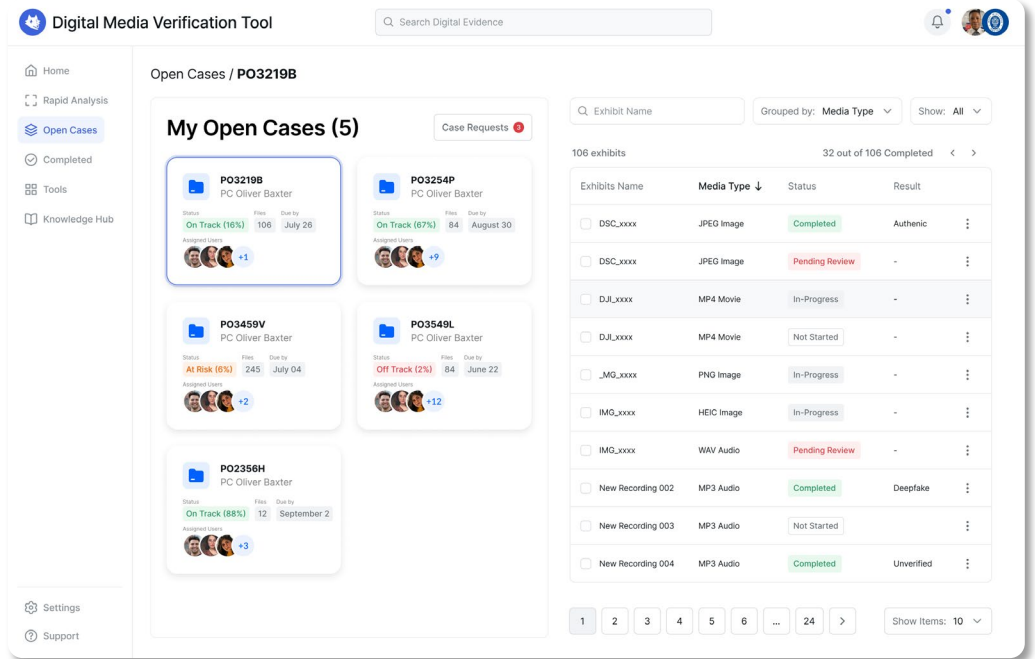
# Digital Media Verification Tool: Conceptual high-fidelity wireframes for evidential police investigations cont.

## Wireframe 03 Verification Case Management

The Case Management Tab streamlines the organisation and prioritisation of open cases, displaying percentage completion and assigned colleagues. Each case folder offers immediate access to all evidence exhibits, presented in a structured table with details such as name, media type, completion status, and results, enhancing the efficiency of the analysis process.



- Used by DMIs to manage their incoming requests for assessments

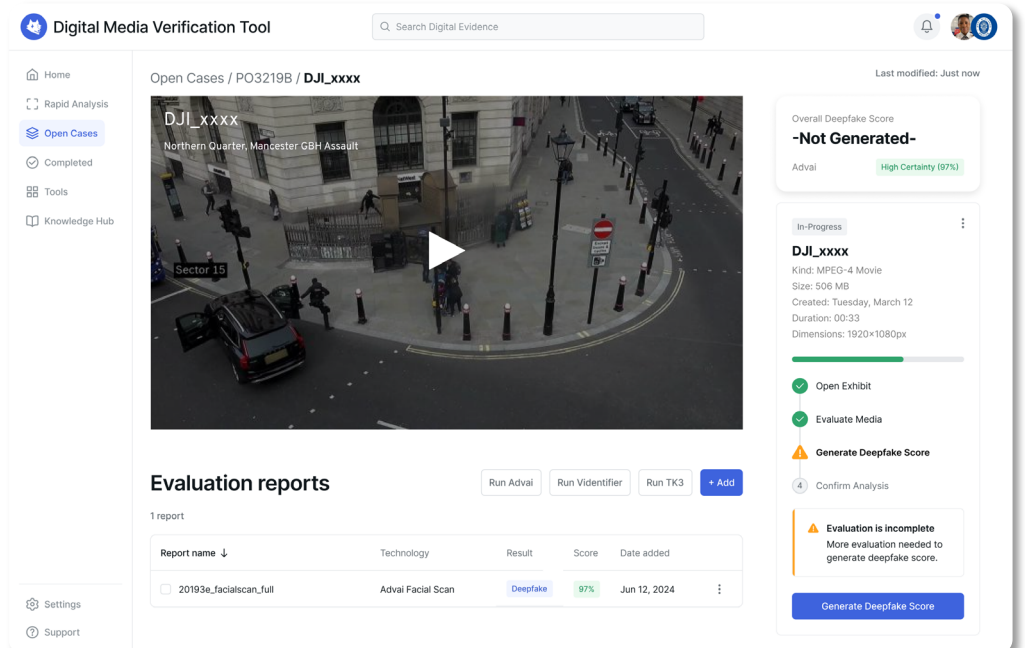


## Wireframe 04 Exhibit Preview and Reporting

The Exhibit Preview and Reporting provides a viewable preview and status tracker for each media exhibit. Quick access buttons launch deepfake analysis tools, with reports stored in an organised table upon completion. The analysis produces an overall deepfake score, along with a qualified confidence rating, for a thorough assessment by the investigator.



- DMIs can view exhibits and select different tools depending on the media type
- Scoring from different tools could be summarised or viewed individually to support decision making





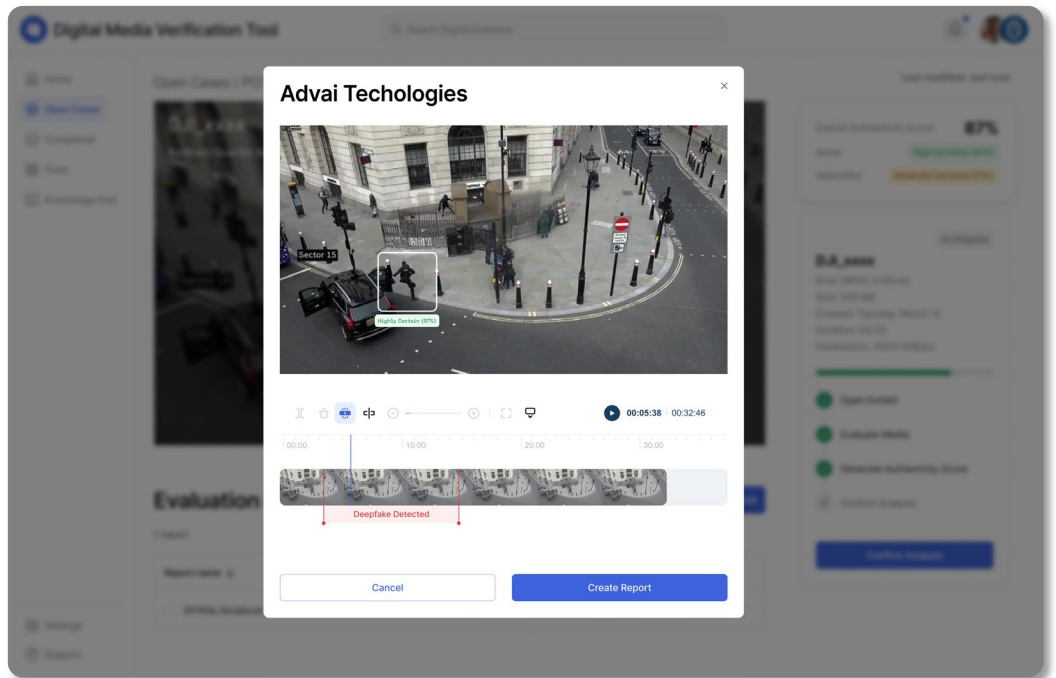
# Digital Media Verification Tool: Conceptual high-fidelity wireframes for evidential police investigations cont.

## Wireframe 05 Digital Media Analysis

A series of tools are used to assess authenticity. Upon initiation, the feature automatically processes the media through these tools to detect anomalies and inconsistencies indicative of manipulation. Classical digital forensic tools are also available to offer an additional layers of refinement for a comprehensive analysis.



- Organisations or departments may choose to pre-select tooling used or allow a DMI to make that decision depending on their approach to risk management
- More advanced analysis would be available to Digital Forensics

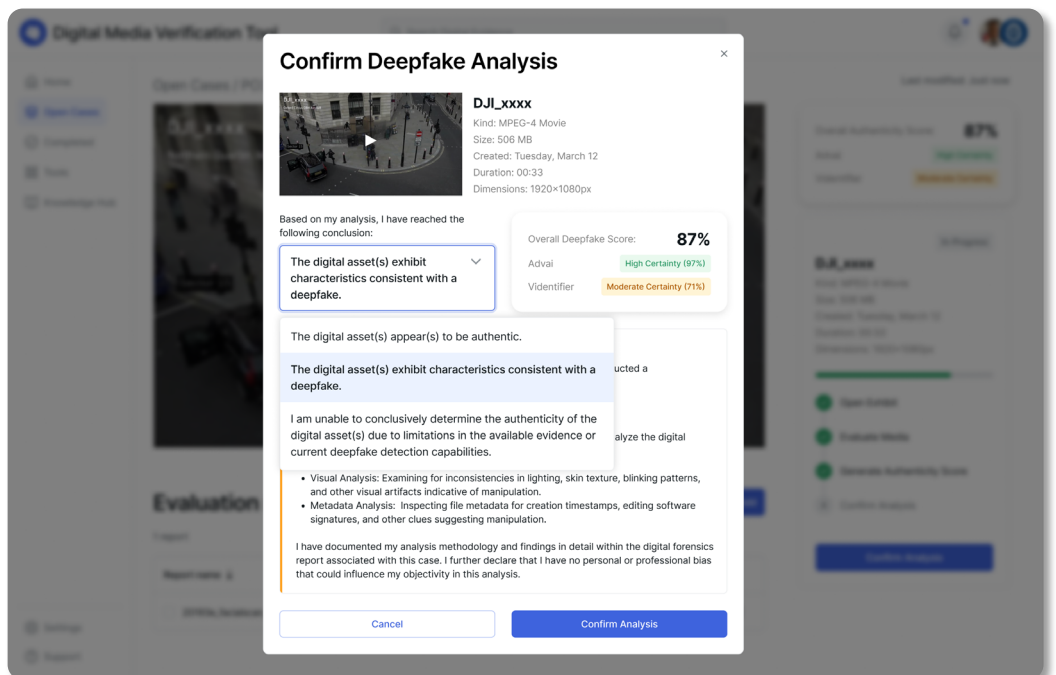


## Wireframe 06 Analysis Confirmation

Once the analysis is completed, a detailed report is generated and appended to the exhibit's case file. This report includes findings from each algorithm used, providing insights into the media's integrity and aiding investigators in making informed decisions regarding the authenticity of the media exhibit based on solid forensic evidence.



- Verification will be a combination of input from tooling and human judgment
- This will create an audit trail of analysis and decision making which will be critical for use evidentially



# Advai have proposed an enabling technical solution that meets requirements across users

The different user profiles have different deepfake detection requirements; necessitating multiple technological needs. Our technical approach provides this nuance, whilst also cohering them in support of the different parts of an evidential chain; up to the need to defend digital media evidence in court.

## Stage 1



**Police  
Constable  
Olivia**

Stage 1 provisions tools such as a filter detector or a binary classifier, to provide frontline officers such as Olivia with an indication of the presence of deep fake imagery.

This enables Olivia to rapidly determine if deepfake indicators, to verify any evidence provided at the scene and progress the investigation. Once the initial analysis is complete, she can upload it for deeper analysis.

## Stage 2



**Digital Media  
Investigator  
Joanna**

Stage 2 provides more a more comprehensive tool to analyse digital media that has been identified by front-line officers.

This enables Joanna who, with specialist training, needs to prioritise evidence for deeper analysis. Joanna can determine if the whole image, or only part of it, is deepfake, and another suspect is identified.

## Stage 3



**Digital Forensics  
Expert  
Michael**

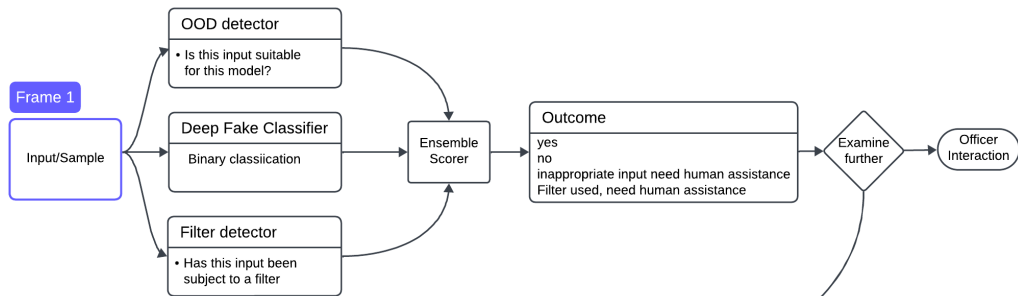
Stage 3 provides support to digital forensics specialists, providing definitive identification of real & fake images, as evidence is finalised for court.

Michael can use techniques such as pixel-by-pixel confidence scores, displayed as interpretable heatmaps, to prove and demonstrate how and where a deepfake has been applied.

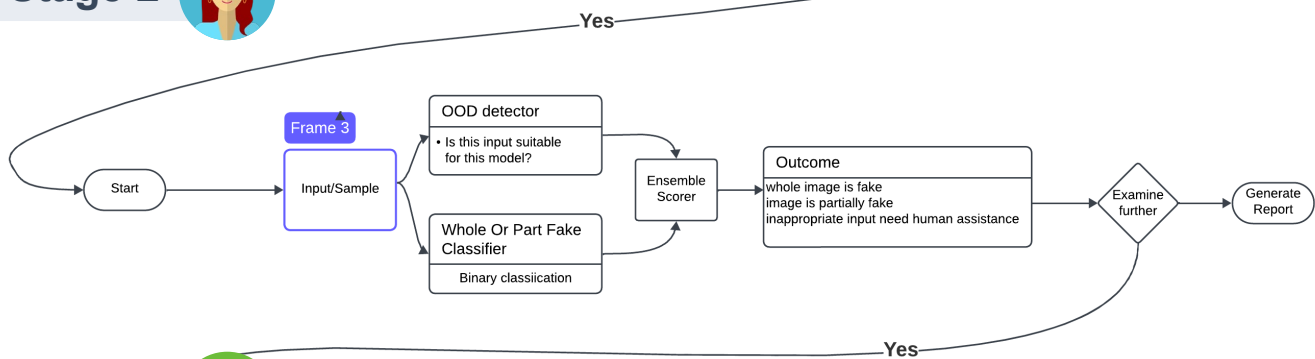


# It is designed to combine strengths of different detectors for effective detection

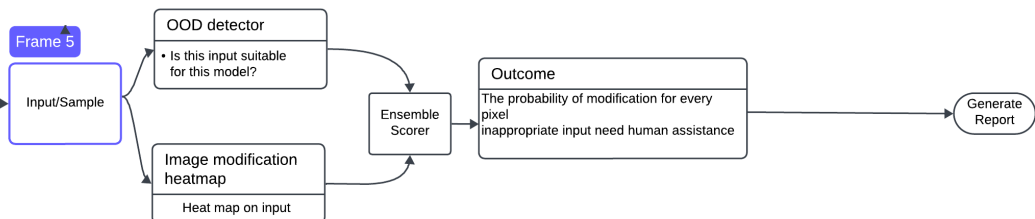
## Stage 1



## Stage 2



## Stage 3



### Key features:

- Each stage designed to be independent, but coherent and in support of the evidential process; as more investigation is required more appropriate tools can be applied.
- All stages have a similar structure and functions as a model ensemble; enabling us to combine the strengths and weaknesses of different detectors. Every model should have an out-of-distribution detector so that new inputs can be assessed for suitability to the ensemble
- They are coherent, and mutually supportive; by progressing images through the stages real images can be gradually discarded and the full details of a deepfake can be ascertained.
- Training, testing, and deploying these pipelines quickly and efficiently will enable us to keep up to date with this rapidly evolving field.



—  
**Bringing  
Ingenuity  
to Life.**  
—



## Corporate Headquarters

10 Bressenden Place  
London  
SW1E 5DN  
+44 20 7730 9000

### About PA.

We believe in the power of ingenuity to build a positive human future.

As strategies, technologies, and innovation collide, we create opportunity from complexity.

Our diverse teams of experts combine innovative thinking and breakthrough technologies to progress further, faster. Our clients adapt and transform, and together we achieve enduring results.

We are over 4,000 strategists, innovators, designers, consultants, digital experts, scientists, engineers, and technologists. And we have deep expertise in consumer and manufacturing, defence and security, energy and utilities, financial services, government and public services, health and life sciences, and transport.

Our teams operate globally from offices across the UK, Ireland, US, Nordics, and Netherlands.

### PA. Bringing Ingenuity to Life.

This report has been prepared by PA Consulting Group on the basis of information supplied by the client, third parties (if appropriate) and that which is available in the public domain. No representation or warranty is given as to the achievability or reasonableness of future projections or the assumptions underlying them, targets, valuations, opinions, prospects or returns, if any, which have not been independently verified. Except where otherwise indicated, the report speaks as at the date indicated within the report.

[paconsulting.com](https://paconsulting.com)

All rights reserved © PA Knowledge Limited 2024

This report is confidential to the organisation named herein and may not be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical or otherwise, without the prior written permission of PA Consulting Group. In the event that you receive this document in error, you should return it to PA Consulting Group, 10 Bressenden Place, London, SW1E 5DN. PA Consulting Group accepts no liability whatsoever should an unauthorised recipient of this report act on its contents.