

# Deepfake Forensics Survey

## Consent and Project Outline

You are being invited to take part in a research project. 1. What is the purpose of the project? The research is exploring the impact and implications of deepfakes/fabricated media in digital forensic investigations and the cyber domain. We aim to gather useful insight and statistics that can aid researchers and practitioners in evidencing this growing issue. We are conducting the research to better understand the needs of forensic practitioners and cyber experts, and to collect data on this emerging issue within their domains. 2. Why have I been invited to take part? You have been identified as a potential participant by the research team or you have scanned a QR code to take part. 3. Do I have to take part? No – it's up to you. You can ask questions about the project before deciding whether to take part. If you do not want to take part that is OK. If you wish to take part: • You will select 'Yes' when asked if you would like to participate. • Submitting the completed questionnaire implies your consent to take part in this project. During the completion of the survey, you can withdraw at any time. You can withdraw by closing the window and/or simply not submitting your responses on the last page of the survey. Partially completed surveys will not be retained. After your response is submitted, it cannot be subsequently withdrawn, as your response is anonymous. 4. What will happen to me if I take part? The questionnaire will be conducted online (using your phone, tablet, laptop or PC) and includes questions that will ask you about the rise in deepfake media in digital forensics and cybersecurity. The questionnaire should take 15 minutes to complete. The responses given to the survey are anonymous, with no uniquely identifiable information recorded. The information participation sheet can be downloaded here: <https://tinyurl.com/3nsfe9jw>

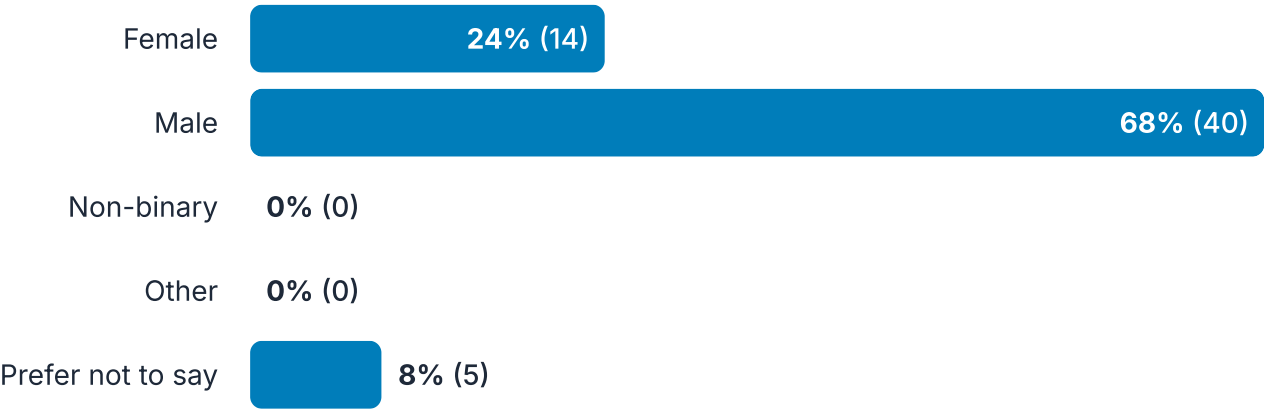
1. Do you wish to participate?By selecting 'Yes' you are confirming you have read the information sheet provided and are happy to participate. You understand that by completing and returning this questionnaire you are consenting to be part of this study and for your data to be used as described in the information sheet provided.

Responses: 59



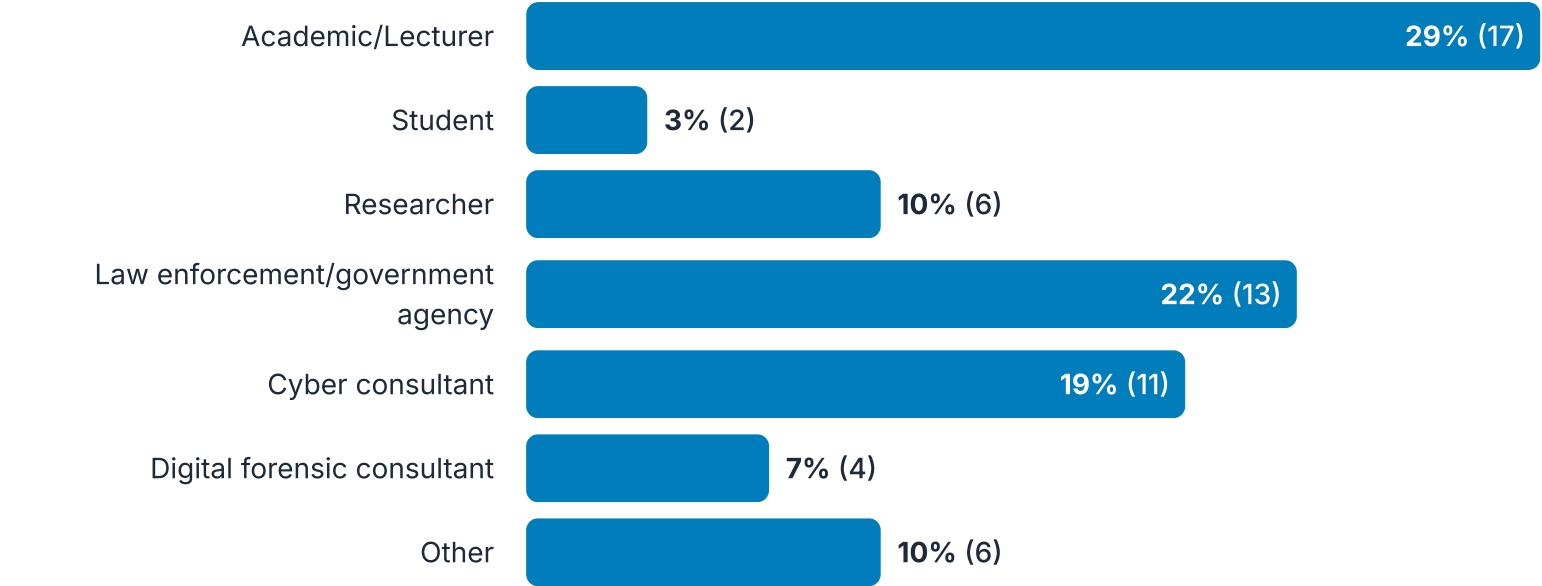
2. Which of the following best describes your gender?

Responses: 59



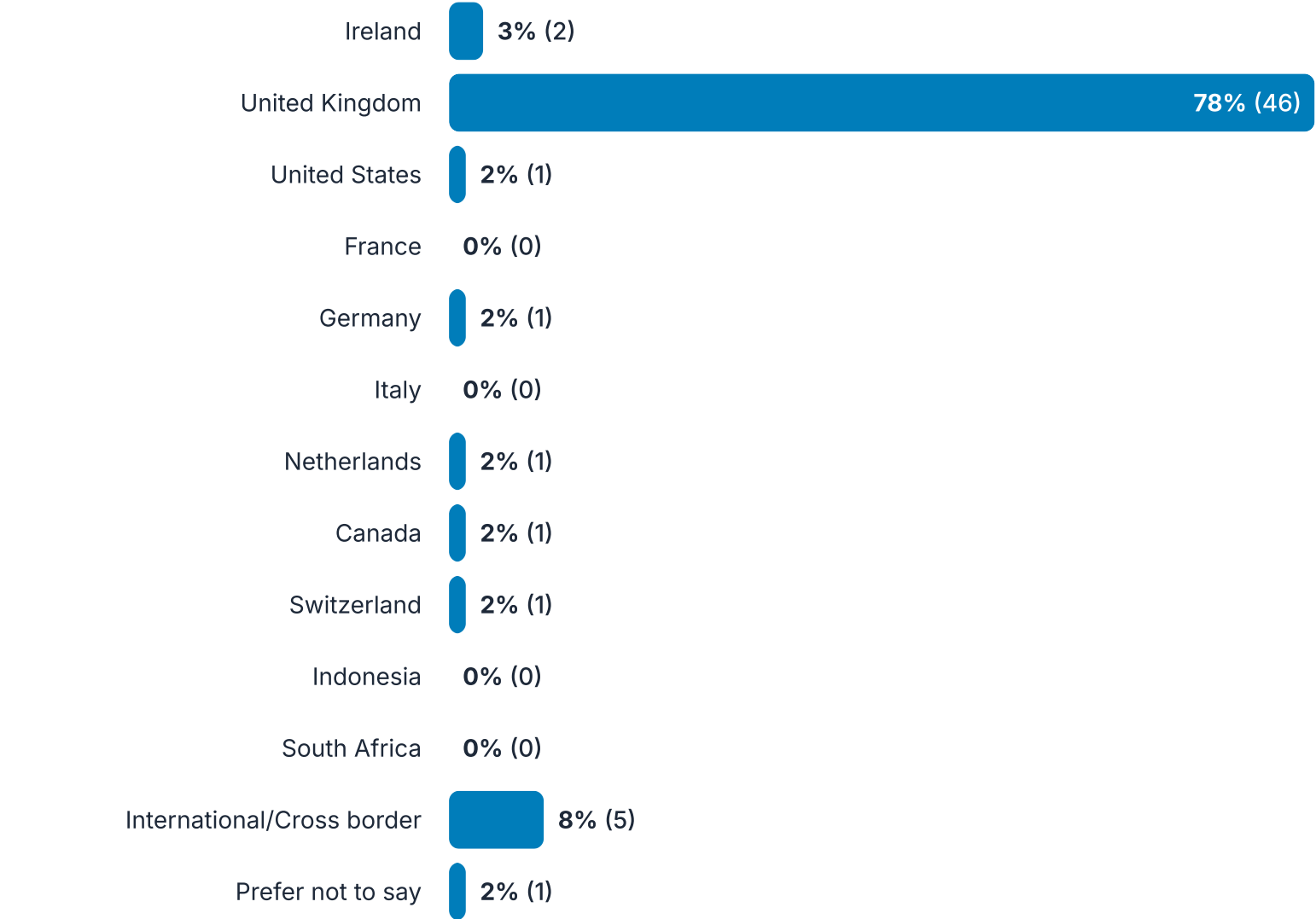
3. What is your current role?

Responses: 59



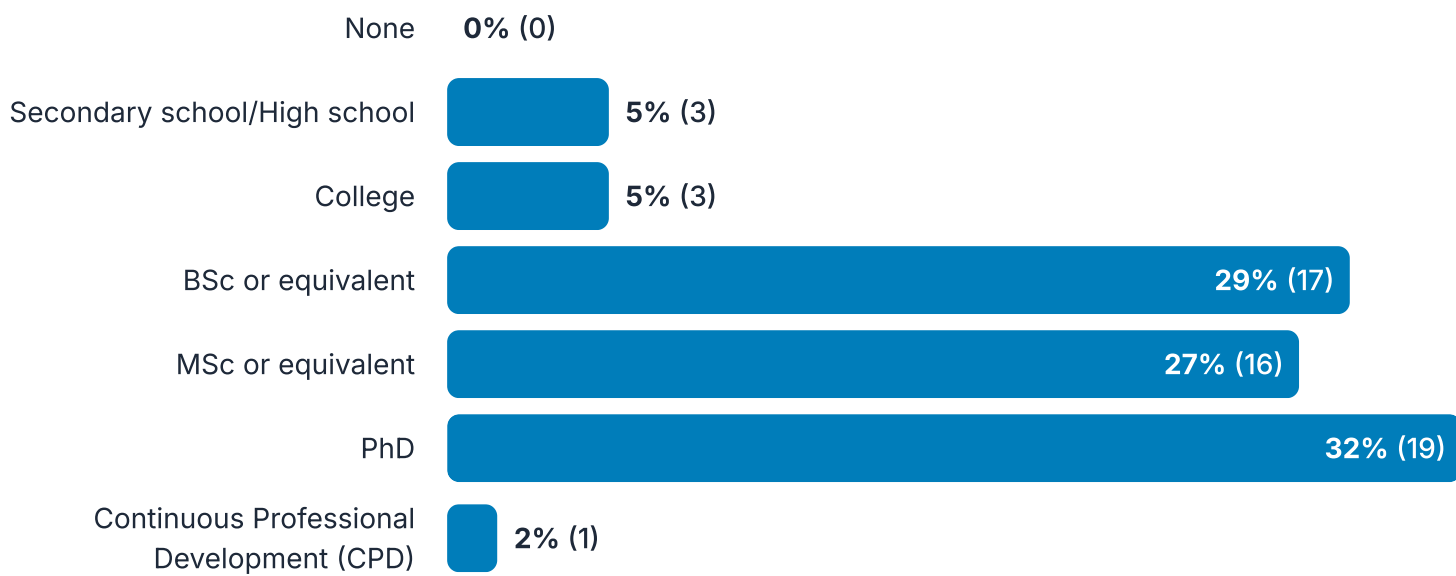
4. What country do you mainly work and reside in?

Responses: 59



5. What is your highest level of qualification?

Responses: 59



## Consent and Project Outline - Duplicate

You are being invited to take part in a research project. 1. What is the purpose of the project? The research is exploring the impact and implications of deepfakes/fabricated media in digital forensic investigations and the cyber domain. We aim to gather useful insight and statistics that can aid researchers and practitioners in evidencing this growing issue. We are conducting the research to better understand the needs of forensic practitioners and cyber experts, and to collect data on this emerging issue within their domains. 2. Why have I been invited to take part? You have been identified as a potential participant by the research team or you have scanned a QR code to take part. 3. Do I have to take part? No – it's up to you. You can ask questions about the project before deciding whether to take part. If you do not want to take part that is OK. If you wish to take part: • You will select 'Yes' when asked if you would like to participate. • Submitting the completed questionnaire implies your consent to take part in this project. During the completion of the survey, you can withdraw at any time. You can withdraw by closing the window and/or simply not submitting your responses on the last page of the survey. Partially completed surveys will not be retained. After your response is submitted, it cannot be subsequently withdrawn, as your response is anonymous. 4. What will happen to me if I take part? The questionnaire will be conducted online (using your phone, tablet, laptop or PC) and includes questions that will ask you about the rise in deepfake media in digital forensics and cybersecurity. The questionnaire should take 10 minutes to complete. The responses given to the survey are anonymous, with no uniquely identifiable information recorded.

6. Do you wish to participate? The participation information sheet can be downloaded here for your records: <https://tinyurl.com/3nsfe9jw> By selecting 'Yes' you are confirming you have read the information sheet provided and are happy to participate. You understand that by completing and returning this questionnaire you are consenting to be part of this study and for your data to be used as described in the information sheet provided.

Responses: 58

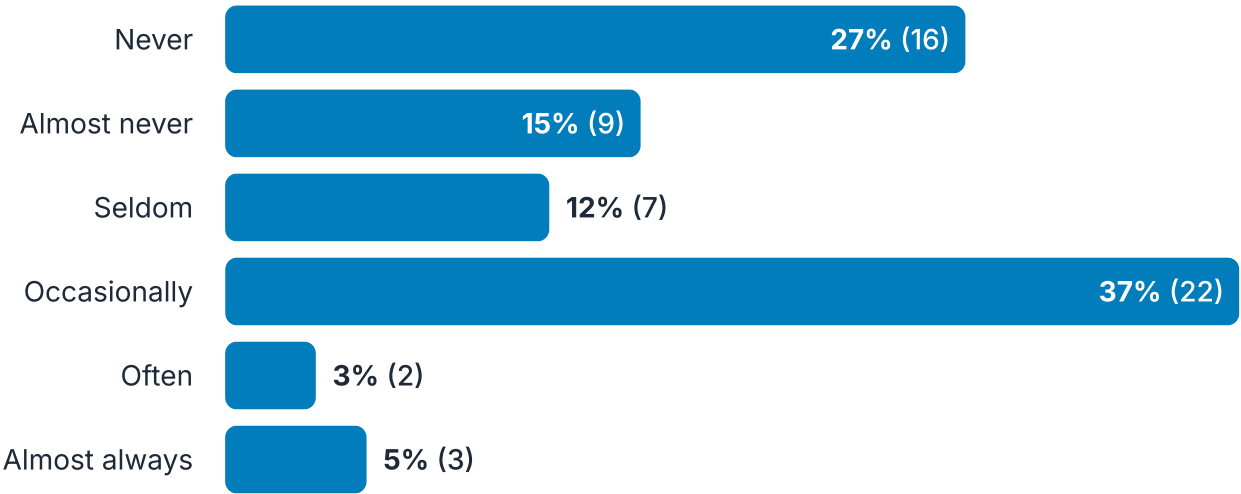


Exploring the impact and implications of deepfakes/fabricated media in digital forensic investigations and the cyber domain.

7. Do you think the rise in deepfakes is an increasing problem for the digital forensics and cybersecurity field? Responses: 59



8. Have you encountered deepfakes (AI-generated forgeries - false images, audio, or images - that appear convincingly genuine) or other forms of manipulated media in your role? Responses: 59



9. If yes, please give an example of how deepfakes could be a problem in your job role: Responses: 35

- Deepfakes image/face swaps of people. Also examples of Deepfake songs made by students
- White collar crime
- I don't see video or image generation posing many problems for education/academia, but the field of digital forensics (my subject matter) will certainly have issues. Media provenance is going to be a huge issue, and, oddly, it's even difficult to inform students about how to approach it at the moment.
- Not necessarily as a problem in my job role, but I have encountered deepfakes.

Fake photos or satire video

Creating a false narrative and increasing time needed to detect them

I previously worked in law enforcement whereby deepfakes were seen often - this included cases like revenge pornography and child sexual abuse cases - I have seen deepfakes used both to intimidate individuals by the act of blackmail, and for personal use. In my current role, in DFIR consultancy, I have not come across deepfakes, but I have conducted some research into deepfakes and how dangerous they can be.

The existence of deepfakes, and increasing public awareness of their existence, increase the potential for jurors to doubt the veracity of any material which could be generated by "AI" systems, potentially making "reasonable doubt" an easier hurdle to clear.

Deepfakes could become a problem when completing victim identification in regards to if someone's face has been put onto rude and obscene photos

It impacts our customers who think we contact them. Deep fake AI generated calls are used in fraud, money laundering etc

Notably the use of deepfakes on social media platforms spreading misinformation and from a professional standpoint how deepfakes are used by criminals

Impersonation leading towards social engineering via password resets etc.

Concerns over integrity of files and data. Questions from above on how we are dealing with it.

Social engineering Video / Voice impersonation Spear phishing Identity fraud Disinformation

Customers pretending to someone they are not, to get services and then never pay.

Often encounter low quality generated forgeries, however there has certainly been an increase in higher quality over recent years. This could present challenges with identifying real victims as part of investigation with generated media becoming increasingly prevalent. Furthermore, it could become challenging to rely upon multi-media in certain cases if unable to distinguish between real and generated content. Especially if AI-generated material fully mimics all aspects of multi-media such as comprehensive metadata (Capture times, geo data, etc.)

Hard to spot, additional time required to analyse when backlog already very high, pressure from management to complete work faster is a bad combo for missing these.

Students will see deep fake images as real and try to use them as part of their assessment. Deep fakes causes issues with authentication of users.

Deepfakes of indecent images of children is starting to become more prevalent with suspects running their own AI models locally and creating images using this method

No but concerns re access and authentication risks by way of spoofing and/or social engineering once deepfake exploits become more common.

AI generated video of know female in a sexual scene

Things like AI generated research reports

With the use of AI deepfakes, they can be used to manipulate the public to sway a certain direction. This is seen across political spectrums especially.

Deepfakes pose a problem for cybersecurity operators because they can be used to impersonate trusted individuals, spreading misinformation or tricking users into revealing sensitive information.

Misinformation and Manipulation

Various media is required for assessments (e.g. screenshots, video evidence of undertaking a task), which can either be manipulated or generated. This makes it difficult to set and mark assignments that are truly assessing a student's level of skill/understanding.

Manufactured / AI-generated work submitted by students. Faked credentials, etc.

Phishing and Spearphishing are main concerns right now.

The attack on the 'human' layer of security will be most prevalent through various forms of Phishing. I have observed Vishing attacks utilising boss' voices utilising deep fakes.

Fake student identities Fake CPD beneficiaries Student video presentations Interview details CCTV evidence

I work for a defence company where deepfakes could cause massive issues to the people operating machinery

Not encountered in my role, although very aware of it through personal social media browsing and training.

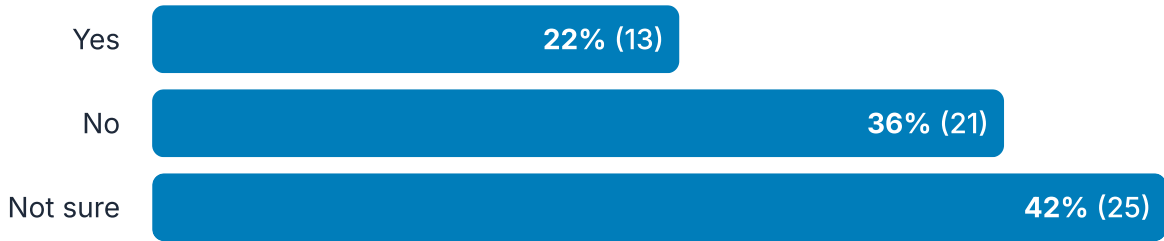
Haven't encountered it yet but it will make my job a lot more difficult once it becomes relevant, issues with proving its a deepfake for court, even realising its a deepfake etc

Targeted phishing campaigns or people committing crimes based on fake media that they have been shown

Deepfake videos or images can be embedded with malicious code or used as bait in: Malvertising campaigns Trojanised media files Social media-based attacks Attackers could fabricate videos of employees engaging in misconduct or leaking data, leading to: Internal investigations Distrust among staff Disruption of operations



11. Do you feel confident in your ability to detect a deepfake? Responses: 59



12. Approximately how many cases/incidents have you been involved in that required analysis of deepfake media? [If valid, how easy was it to determine authenticity of media?]. Responses: 49

- I have been exposed to a few with varying successes. The tools gave different results.
- 0
- 2. Easy in our case due to inconsistent metadata and impossible shadows
- I don't do case work.
- N/A
- N/A but half of the time I am not sure
- Minimal
- Cases - around 10, these would include deepfakes that range from poor and easy to spot, to sophisticated and hard to identify, in which case further forensic analysis would be required.
- No idea!
- 0
- None at this moment in time

2, it was not easy

Minimal, these have been more during research as I do not actively investigate anymore.

Two, however they were both easy to detect via visual inspection (missing teeth, multiple fingers, facial hair errors etc)

0

None to date

1-2 per week

None

I haven't work on cases specfically focused on deepfake media, however I have worked on cases involving image categorisation. During which I have encountered many fake/generated multimedia. Generally speaking many were low quality and easily identifiable through notable features such as image distortion or poorly blended images. That being said, their has been a increase in more convincing fake media likely due to the raise of AI powered tools available online. I imagine this could become a greater issue as better AI models are trained and made publicly available.

0

Hard to say, not many but increasing. I'd say 10 a year currently. Some more when the case does not involve it specifically but they were found during investigation. When I started in 2019 I'd say I had 1 job.

Not relevant to job role.

Three Yes it was easy to detect but only because I know what to look for

None

None

zero

None as of yet

I have encountered 1 case, some images wheee convincing however others where clearly AI generated

None yet.

0



none

Social media plays a large part in spreading deep fake material. I have seen many instances of this especially with Ai generated videos depicting famous celebrities or politicians in saying something that they never have.

2

na

0

None

0

No deepfake media incidents, as required media is very specific and obscure.

N/A

None

None

This isn't a ever day occurace however the numbers are increasing.

0

3-4

None personally just heard of them

N/A in my role.

None

No specific ones but many of the ones I have seem would be obvious to the naked eye

Look for visual clues, check file metadata for signs of editing or inconsistencies in timestamps - perform media forensics as it is basically the same domain

13. In your opinion, how effective are current AI-driven detection tools in identifying deepfake content?

Responses: 51

- It is average and can vary.
- Not effective
- Somewhat
- 50/50 good to detect, unreliable for evaluation purposes
- From a literature point of view, they're only good for what they're trained for. Generalisation is a big problem at the moment.
- mediocre, a lot of work still needs to be done in the area.
- not really
- It depends on the tool available
- They are OK
- I have not used AI driven detection tools. Only for research purposes, and some deepfakes were not identified as deepfakes.
- Not very. Detection will always lag production and production is getting better.
- Unknown
- It would be better to personally look at it and identify it
- Not very effective at all
- Tools are only as good as the person using them. The human eye combined with training can be more effective.
- Underdeveloped.
- Not applicable
- Effectiveness can vary and also depends on which tool used
- Improvements could be made

Not very effective.

Not very effective

50/50

Unknown and not used, I've only done self analysis and use of standard tools and exit tools. Some media embeds data making it easier but I imagine I have missed the ones which are harder to detect. Eg not having 6 fingers etc

The current tools give too many false positives and false negatives.

Not useful at all

Great

I think they are good, but I think AI is moving faster than the counter measures n

i have not used ai-driven detection tools

Quite good

It all depends on a case by case basis, only select forces are using AI to help spot deep fakes

I haven't used any but would be optimistic of high accuracy [dependent on AI model] but acutely aware of potential limitations in real world deployments with various scenarios to consider.

Not effective

Not sure

Average. Certain deepfakes may be able to still bypass the detection tools.

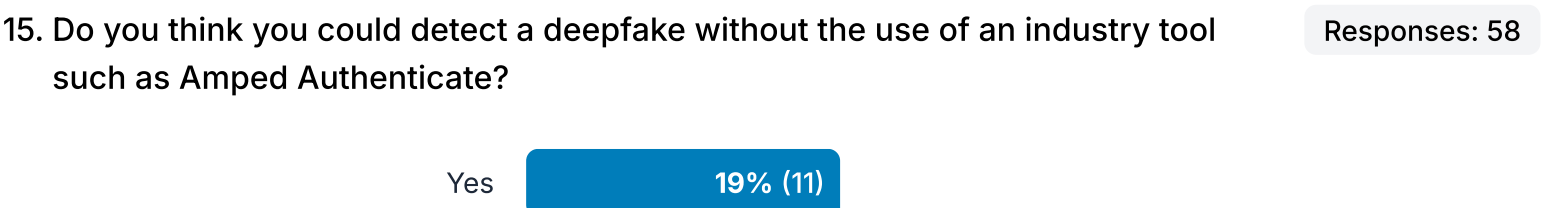
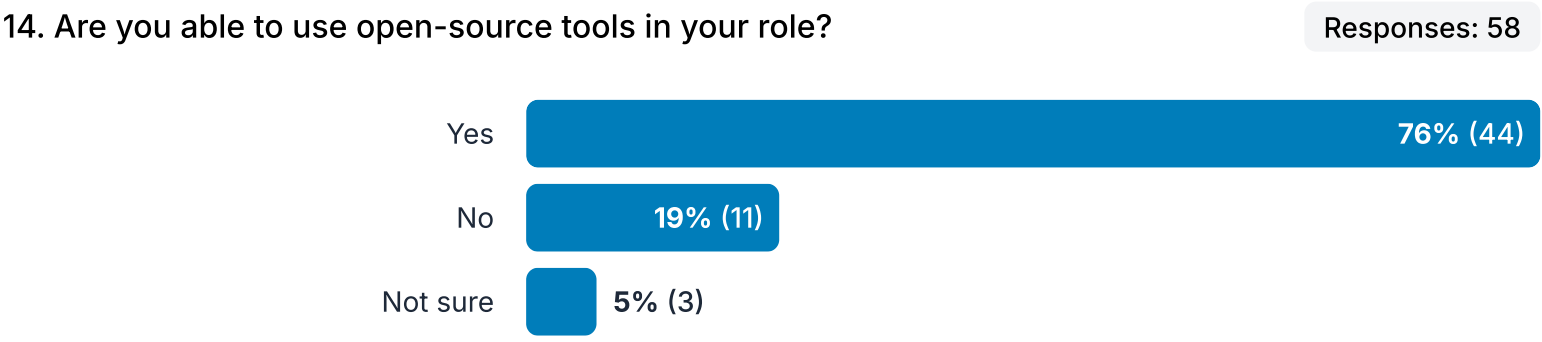
Not good enough

We dont use them or AI to analyse our data. All our analysis happens on a closed loop system and is not connected to the internet. Therefore, unless a software provided has an offline AI package we cannot use the AI functions.

I personally have not checked any tools and not sure how effective they are

I'm not aware of any.

- Current AI-driven detection tools are reasonably effective with certain types of deepfakes, especially those generated by older or less sophisticated methods, but their effectiveness is declining as deepfake technology improves.
- It's a mixed bag but at best they can be described as reasonable
- Mixed bag
- They are okay
- Not at all
- This is a work in progress.
- Largely ineffective
- Variable
- Somewhat
- Unsure
- Not as effective as humans
- Some are good but they can be expensive
- Average



No

19% (11)

Not sure

62% (36)

## 16. What industries or sectors are most vulnerable to deepfake-related threats?

Responses: 51

Society and education

Marketing

Outside of CSAM and other sexual crimes, fake news is the worry as it can be difficult to combat things after the fact. The Scottish government had a call out several months back looking for advice on how to handle the risk of their parliamentary recordings being faked.

At current - Political, government, romance scams

dont know

Mainstream media

Public sector

I would say the industries most vulnerable are finance and law enforcement.

See comment above - but anything involving phishing-type activity can be enhanced by deep-fakery of imagery, video or voice.

Unknown

Victim identification and many cases of cp

Finance!

I would say all industries could suffer due to deepfakes, this could be anything from financial industries to e-commerce.

Finance Health Retail

Digital Forensics, Tech Industry, Banking.

Public sector/finance/education. It depends on the quality of deepfake

All

Smaller SMEs that require verification of identity and are not able to invest in systems that can identify deep fakes.

. Digital Forensics, . Cyber security, . AI driven industry and research.

The elderly Politics

Creative, public people e.g politicians, influencers law enforcement to detect them

News media and politicians. Celebrities and those that would be believed by the general public.

Every industry and area

Academic

I think Civil service sectors and security sectors would be particularly at risk. Online banking possibly

unsure as i have not encountered deepfake in my digital forensic investigations.

Police / cyber crime

Police, the rise in deep fake indecent images of children

All, but in particular healthcare whereby resources are limited, cyber security training isn't prioritised and personnel subject to pre-existing pressure.

Not sure

Social Media

Finance and Banking. Scammers may request funds transfer or access accounts illegally. I see a lot of fake accounts in the e-commerce industry when attempting to purchase tickets. Media and Politics These can be fake recordings of politicians making promises and different things to manipulate the voter base.

Finance, healthcare, governmental, media

I think the risk is more inline with the harm done by deep fakes spread around the web with people taking them as genuine, they are damaging to reputation etc.. a deep fake could be an issue in a host of investigations if not detected at an early stage. Could a deep fake lead to false/bad intelligence and executive action there on?

IIOC, revenge porn,

Industries which in charge of building new AI models, and Cyber Security industries

Politics. National security. Government.

Finance and Banking, politics, media, insurance....

Nearly all industries/sectors are vulnerable to deepfakes but I'd say critical infrastructure, banking/investment and the arts.

Any that relies on remotely-submitted documentation / images. Government, finance, health?

Finance, software, government, infrastructure

Politics, Media

All home working jobs are most susceptible due to means of communication. This won't be industry or sector specific, it will be to the goal of the attackers, I.e. Financial or political gains.

All

All

Almost any industry that relies on video or digital evidence for authentication or validation purposes

Military

Organisations such as Insurance, Healthcare, political.

Politics, entertainment, etc.

There are many examples of the financial sector being targeted, but also political.

All but some examples would include: Government and National Security Disinformation campaigns: Used to spread fake political messages or impersonate officials. Diplomatic sabotage: Fabricated videos of leaders making inflammatory statements. Cyber-espionage: Deepfake voice or video used to gain access to secure communications. Corporate and Financial Services CEO fraud: Deepfake audio/video used to impersonate executives and authorize fraudulent transactions. Stock manipulation: Fake news or videos affecting investor confidence. Insider threats: Fabricated evidence used to discredit employees or manipulate internal investigations. Retail and E-commerce Customer service fraud: Impersonation of staff or customers to exploit systems. Fake reviews and endorsements: Deepfakes used to simulate influencers or satisfied customers. Brand damage: Fabricated videos of product failures or scandals. Media and Entertainment Fake news: Deepfakes used to create false reports or interviews. Celebrity impersonation: Used for scams or reputational harm. Content piracy: Manipulated media used to bypass copyright protections.

17. What advice would you give to forensic professionals on dealing with deepfake-related cases?

Responses: 49

Keep reading research and domain news. They are getting harder to detect.

Manual analysis and metadata are going to be more reliable, but time consuming. Ideally they need training, but that will only be valid in short bursts as things advance quickly.

Read up on the literature surrounding it. Make sure you're educated in how and where it can be present, and reach out or ask if unsure.

dont know

Publicize the threat more

Share more knowledge on it

Use detection tools cautiously - don't take anything as gospel, review in detail, perform real deep forensic analysis where required.

Panic!

Unknown

Be careful with what looks to be real it could potentially be deepfaked

Use any tools that are available, research the topic, speak to others to exchange ideas and threats

Don't believe everything you see, question everything.

Work on other aspects of security which don't rely on a single factor which can be affected by deepfakes.

N/A

Keep up to date with literature and keep software updated

Use detection tools Verify file origins, timestamps for authentication Cross referenced evidence Educate end users

N/A

It is important to keep upto date with the latest issues arising in tech, espically in how it would impact out ability to rely upon artifacts recovered during investigations or new challenges we face as a result of notable



developments. It is important to encourage education/awareness of current and future practitioners, encourage regulation where applicable, and support research into countermeasures.

Try multiple techniques. Without an asserted original it's difficult to determine if it's been changed

Naturally we are all inquisitive so just continue to be that way and not to blindly trust one source. Use multiple tools to review media

Try to use the current tools, but keep an open mind to their accuracy. It may not be possible to authenticate an image or video.

Learn to think about evidence being deepfake and have it tested

verify to ensure if it's deepfake or original

Taking time to examine the image using all tools available and seeking second opinions

unsure as I have not encountered deepfake in my digital forensic investigations.

Always look for any clues that might reveal that it is a fake, such as, discoloring or the image / video is pixelated

Personally I in my role at the moment would not be able to give any advice being in a placement year.

Clear evidence handling and transparent technical examination

Metadata is the key for knowing the truth

Education is key - the public/students need to be better educated on how to utilise critical thinking, recognise their own bias and be on the lookout for agenda (this will all help in better spotting misinformation or false data, e.g. deepfakes).

Educate clients on AI related risks and providing expert testimonies that will stand up in court.

Stay updated on detection tools and research

Be sure, and be certain and know the legislation around them and possible offences

Not sure

I would recommend trying different technologies to cover AI generated deepfakes instead of AI itself also

Unfortunately, due to the lack of exposure with deepfakes, I don't have any advice.

Using multi-layered detection approach. Document everything, stay up-to date, collaborate.

To remember that there can often be many unseen victims of deepfake media. Models generating the content have to be trained on viable source material. The provenance of this source material can be questionable or illegal.

Remember the basics, artefacts in the metadata might yield some clues. Look for inconsistencies and rely on common sense in the first instance. Consider the context and environment and identify any points that stand out. Consider layered or multi-modal approaches that utilise multiple data points/sources.

I would say to just keep up to date with tooling and training

By the time a the incident has reached this level, the deep fake will be an article or pivot point for the analyst/ victim still had the deep fake it will give context to the request, i.e. taking the victim to a c2 server or request for information ect.

It is an issue that needs to be taken more seriously, and urgently

Research

Trust your gut instinct

Be cautious and meticulous

Look for the common signs that would help in identifying a deep fake such as looking at hands, fingers and even thinking about the context of the photo / video.

I haven't personally dealt with deepfake-related cases yet, so I don't have any advice.

Keep discussing the topic and share findings and best practices

Forensic professionals should stay current with deepfake generation and detection techniques, use specialized tools to analyze synthetic media, and rigorously preserve digital evidence. They must also collaborate with legal teams, educate stakeholders, and integrate deepfake analysis into broader cybersecurity and incident response strategies.

---

18. Is there anything else you would like to add?

Responses: 30

Is there an industry approved tool or approach? There isn't much literature on what to do but many say it is a growing problem.

This is an interesting topic

Digital forensics field and DFUs need to be unified in their approach and get a handle on it before it gets exploited

N/A

We also need to worry about the propensity of AI to give false results in detection and processing of other material. Until we actually have explainable AI which can stand up to human scrutiny and confirmation of results, any evidence touched by AI can be used of being fake itself.

No

N/a

Finance, loans, banks are the primary targets these days.

Deepfakes extend further than just audio visual, text based fakes have been around for a while also.

No.

N/A

I dont trust detectors due to the technology moving so quickly and the diversity of generation methodology. Difficult to handle

No.

no

No

no

Just that the problem spans multiple industries/fields, and is a serious one!

I appreciate you taking on this research project. It is immensely important and will only get more common in all industries including mine in e-commerce. It is vital that we stay on top of the growing threat to AI and Deepfake technology.

Cross verify sources when ever possible

no thanks

There probably is a gap in the market on deep fakes, giving CPD to police forces would be beneficial as the world of digital is forever changing and trying to keep up with the emerging threats etc is sometimes difficult with our workloads.

No

It is sad to see the human impact and the tragic outcomes that the weaponisation of such tools can have. Attacks can often feel a lot more personal, isolating and overwhelming (especially in cases of deepfake fraud, indecent media or porn etc.). This problem is only going to get worse and the generated media will become more difficult for the average person to identify.

No

This will become another interesting vector that becomes very interesting when used in spear phishing when communication methods match. I.e a boss who regularly signs off payments via a phone call with time constraints and high pressure environments. This could be a large area to exploit. Although this will grow and more technical controls can be added. Similarly to traditional email attacks, educating staff will be the biggest form of protection.

None

No

No

I would be interested in seeing the follow up report from this survey

No