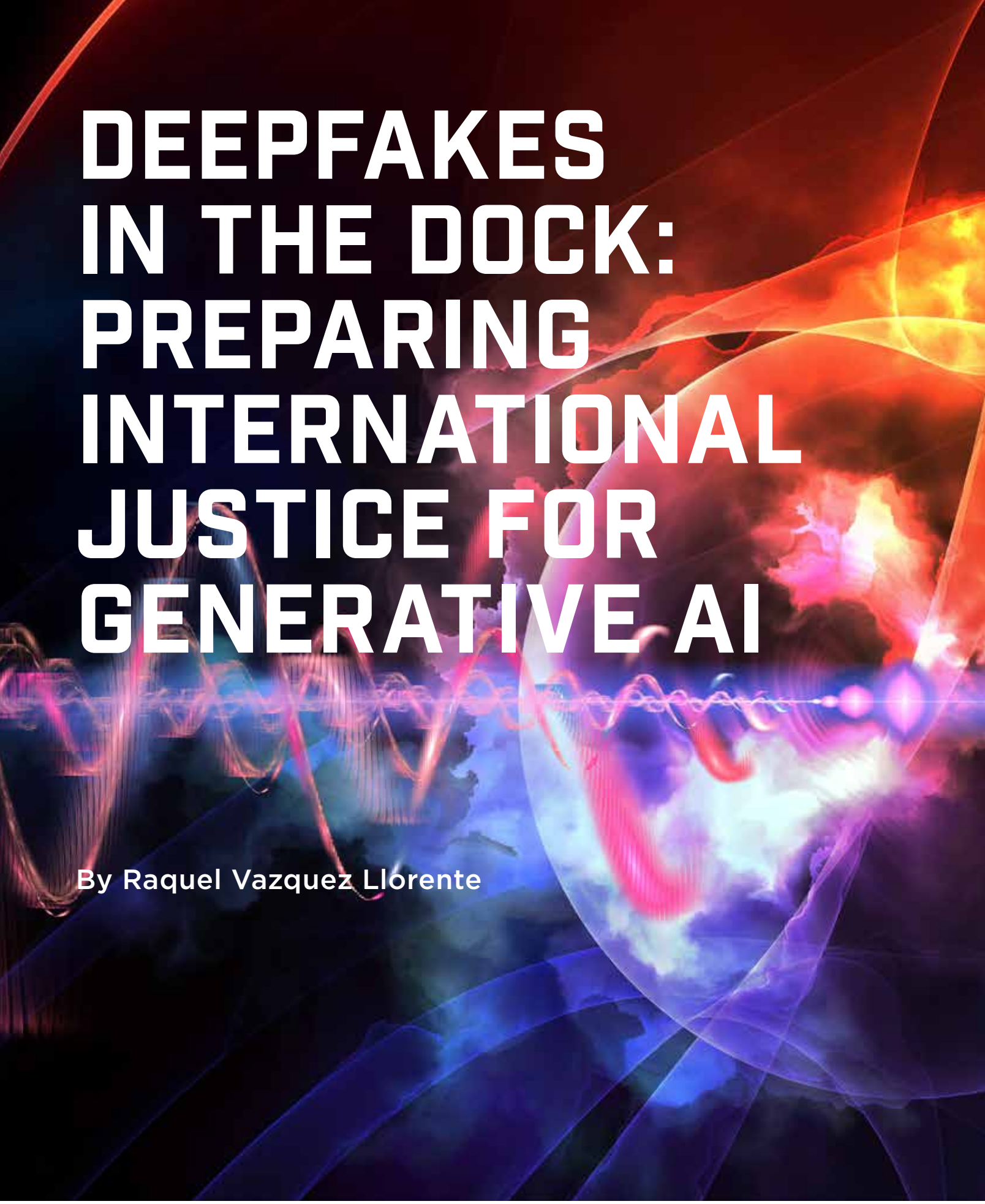


DEEPFAKES IN THE DOCK: PREPARING INTERNATIONAL JUSTICE FOR GENERATIVE AI

By Raquel Vazquez Llorente

The history of audiovisual media is the history of audiovisual manipulation. The first camera came into the world in 1816, although it was not until 1888 when American businessman George Eastman started marketing a device under the name of "Kodak." Cameras were commercialized more widely in the early twentieth century, but even before the average consumer was able to take their own photographs, the world had already seen its first trial for audiovisual manipulation back in 1869. William Mumler, a jewelry engraver from Boston, took a selfie (or, back then, a "self-portrait") that revealed the shape of his deceased cousin on the image. As the story goes, initially Mumler shared the ghostly photograph with a friend as a joke, but seeing the amazement of his colleague, he thought he could make a lucrative business out of "spirit photography" by taking images of people and conjuring their loved ones to appear on camera. These manipulations seem to have been the result of nondivine intervention, showing us an early example of double exposure, by which a previous image made its way into another photograph that used the same glass plate for producing the negative. After a few years of cashing in on the grief that the American Civil War had brought, he was accused of fraud. During the trial, another photographer testified as a witness, having produced himself a fabricated image of a client with the "ghost" of Abraham Lincoln to demonstrate the manipulation technique. Mumler was eventually acquitted.[1]

The Mumler trial not only represents one of the earliest recorded cases of malicious audiovisual manipulation, but it also highlights two powerful dynamics that carry throughout history. First, many of us are willing to place trust in what we see as long as it aligns with our worldview. Second, even when demonstrating the ease of forgery, it may still be difficult to prove in court that any editing took place on a given image. Advancements in audiovisual manipulation, specifically deepfake technology, combined with the commercialization of artificial intelligence (AI) tools that allow anyone with an internet connection to create realistic synthetic images or audio, are now bringing novel challenges to the courtroom. While the impact of generative AI and synthetic media in the information landscape has been the subject of heightened scrutiny since shortly after ChatGPT was made publicly available in November 2022, the effects that deepfake technology can have on international criminal justice are yet to be discussed with similar intensity. This article aims to start a deeper, more informed conversation about how to prepare international courts and tribunals, especially the International Criminal Court (ICC or the Court), for developments in the field of digital evidence and media synthesis.

DEMYSTIFYING THE TECH: GANS VS. DIFFUSION MODELS

Deepfakes, a portmanteau of "deep learning" and "fake," have evolved rapidly since they first appeared in 2017 in a Reddit group that was trading content of celebrities whose faces had been swapped into videos (including pornographic movies). Deepfake technology harnesses AI to create realistic images, videos, and audio recordings that can oftentimes be indistinguishable from reality. Some of the most common techniques leverage Generative Adversarial Networks (GANs) and diffusion models. GANs involve two neural networks that are pitted against each other. One is the generator that creates the examples; the other one is the discriminator that evaluates whether they are real or fake. With each iteration, the generator refines the output.[2] A GAN trained on images of faces can produce new synthetic faces that look realistic.

On the other hand, diffusion models are a type of deep generative model that add noise to the training data and then reconstruct the data by reversing this process. While GANs are like a legal debate where each side sharpens the other's skills, diffusion models are more about taking a broad idea and refining it into something clear and detailed. They are remarkably adept at generating high-quality images that

can sometimes beat the capabilities of GANs. Whereas GANs excel in creating faces and expressions, diffusion models can craft detailed and realistic textures and patterns, making them a powerful technique to fabricate convincing environments or contexts. Diffusion models are responsible for much of the progress over the past two years in the image domain, and they are behind tools now known to the public, such as Dall-E or Stable Diffusion.

THE IMPACT OF GENERATIVE AI ON INTERNATIONAL JUSTICE

Deepfakes, or claims of AI manipulation, have begun to surface in countries embroiled in armed conflicts, experiencing mass violence, or under the thumb of authoritarian regimes. In late 2018, a video of President Ali Bongo of Gabon, released by his office partly to counter claims of his ill health, was dismissed as a deepfake by the opposition, precipitating an attempted military coup.[3] Experts have not been able to conclude unanimously whether the video is the result of AI manipulation (although it was probably not a deepfake). In Myanmar in 2021, Phyo Min Thein, former Chief Minister of Yangon Region, confessed on camera to having paid gold and cash to Aung San Suu Kyi as part of a corruption scheme. The video was likely to have been produced under duress rather than by AI intervention.[4]

In Venezuela and Burkina Faso, governments or their sympathizers have misused commercial software intended to be for company training and product demos, to produce favorable propaganda.[5] In Mexico, candidates in the 2024 electoral cycle have seen themselves or their voices deepfaked into videos or audios.[6] More notoriously, in Ukraine, deepfaked President Zelenskyy called the troops to surrender in a video rapidly debunked by the government.[7] In Sudan, "leaked recordings" of Omar Al Bashir, the former leader who has not been seen in public for a year, were suspected of being manipulated.[8] The organization I work for, WITNESS, has also received tips of two audio files and a video circulated on official social media channels of the Rapid Support Forces (RSF) or of their leader, Mohamed Hamdan Dagalo, who is believed to be dead. We consulted AI media forensic experts, and they all concluded that there did not seem to be indications of AI manipulation—with the caveat that models do not perform to the same level of accuracy for non-English languages. Neither does this mean the absence of other forms of manipulation, such as more traditional editing. In the recent conflict between Palestine and Israel, synthetic media portrays Israeli refugee camps,[9] crowds marching in support of Israel,[10] military attacks taken from video games,[11] and children in need.[12]

These are just a few examples of how AI is creating confusion and mistrust in what we see or hear. As deepfake technology grows more accessible, the volume of synthetic media is increasing, swamping the information ecosystem and impacting most countries, including those under the jurisdiction of the ICC. Gabonese officials and opposition members were once under preliminary examination for crimes against humanity and incitement to genocide.[13] The situation in Myanmar/Bangladesh is under investigation.[14] The prosecutor of the ICC resumed in 2023 their investigation into the situation in Venezuela (for transparency, I was one of the lawyers involved in the communication submitted in 2022 by the Clooney Foundation for Justice and Foro Penal to contribute to the the ICC's investigation).[15] While the Court has yet to act on the allegations of crimes against humanity in Mexico, victims' groups and civil society have been calling for an investigation into mass disappearances for nearly a decade.[16] The ICC has issued arrest warrants for war crimes in Ukraine against Vladimir Putin, president of the Russian Federation, and Maria Lvova-Belova, commissioner for Children's Rights in the Office of the President of the Russian Federation.[17] Al Bashir was the first sitting head of state with a warrant of arrest by the ICC and the first person to be charged by the Court for the crime of genocide.[18] The prosecution is also investigating alleged crimes by the RSF.[19] According to a pre-trial Chamber decision from 2021, the ICC has jurisdiction over Rome Statute crimes in Gaza and the West Bank, including East Jerusalem, and over crimes committed by Palestinian nationals or the nationals of any state parties on Israeli territory.[20] For the past two years, the Office of the Prosecutor has been "steadily increasing the resources and personnel for the Palestine investigation."[21]

Looking at the countries listed above, it is not a question of whether an institution like the ICC will face a deepfake problem, but how badly it will impact justice. U.S. Ambassador-at-Large for Global Criminal Justice Beth Van Schaak has highlighted the urgency of the situation: "Now we have the ability to use synthetic media, generative AI, [. . .] and so you can imagine creating pieces of evidence that could actually infect a legal process."[22] While the proliferation of deepfakes risks false information being accepted as true, the larger threat is that the possibility of AI manipulation is making it easier for the public or those in power to dismiss real content as fake. If this general mistrust creeps into the courtroom, it may poison the well of all audiovisual digital evidence.

The ascent of deepfakes raises the specter of doubt, turning audiovisual content from important probative material into potential red herrings. While deepfakes could be used to frame innocent parties or for creating "evidence" of atrocities that never occurred, looking back in history and at the jurisprudence at the ICC,[23] synthetic media will most likely have two effects on judicial processes for core international crimes if left unaddressed. First, the burgeoning volume of deepfakes will overwhelm the capacity for analysis of institutions in charge of justice and accountability. The conflict in Syria already sounded alarm bells and catalyzed the application of object recognition and frame analysis technology to images circulated mostly online; however, traditional verification methods are no longer sufficient. Second, the leap forward in deepfake technology will lead

to acquittals of those guilty of crimes by inserting doubt on authentic evidence.

FUTURE-PROOFING TRIAGE AND ANALYSIS AT THE INTERNATIONAL CRIMINAL COURT

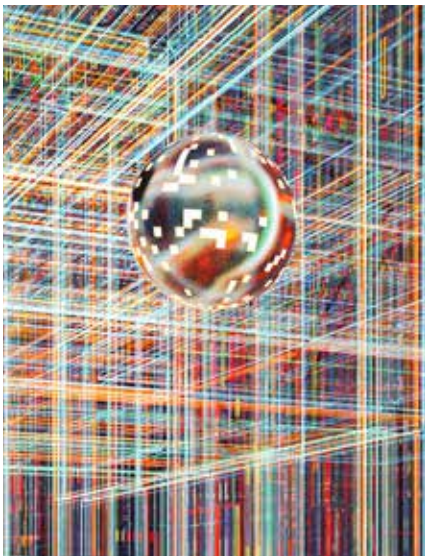
The ICC prosecutor has the obligation to investigate both incriminating and exonerating circumstances equally.[24] Hence, sifting through the avalanche of content to identify authentic evidence is critical. In an academic article I wrote with Lindsay Freeman, we examined how the volume, velocity, and volatility of digital evidence would increasingly affect trial proceedings at the ICC, and we detailed specific recommendations to help the Court address some of their most immediate needs.[25] With generative AI rapidly reshaping the landscape of truth in armed conflicts and situations of mass violence, differentiating genuine content from manipulated media is now assuming a new urgency. This is not just a technical challenge but a fundamental threat to the course of justice.

Institutional responses must evolve in lockstep with technological innovations. As deepfake technology becomes increasingly advanced and accessible, the ICC can counteract the threats this brings by investing in three areas: provenance infrastructure, detection technology, and analytical and forensic expertise. Given the significant resources required to develop and deploy state-of-the-art technology and the capacity gap on media forensics (not unique to the ICC), these approaches will require multistakeholder collaboration, notably with the private sector. None of these solutions will be sufficient on their own, but combined they can be what tips the balance from impunity to justice.

Provenance Infrastructure

The first line of defense against deepfakes is provenance infrastructure—a suite of technology solutions that can help trace the origin of a piece of content and any modifications since its creation. Provenance tools embed metadata into a piece of media, creating "content credentials" that can help with the verification. The human rights sector has been a pioneer in piloting apps for human rights defenders documenting abuses that automatically collect metadata, for instance, CameraV and Proofmode. In the last few years, we have seen provenance technology gaining traction in the form of standards spearheaded mainly by the Coalition for Content Provenance and Authenticity (C2PA), of which WITNESS is part.

Most of the solutions available can provide helpful markers to verify content but have crucial limitations from



a litigation point of view. For international criminal justice, the most useful approach to provenance is what is known as "controlled capture," specialized software that incorporates data points at the moment of recording an image or audio file and maintains the chain of custody throughout the life of the footage.[26] To date, only one solution has been developed that meets these criteria off the shelf and has successfully been tested in a court of law: the eyeWitness to Atrocities app, incubated at the International Bar Association (in the spirit of full disclosure, I was involved in setting up the organization and was one of the lawyers working on the first case that used content collected with the eyeWitness app).[27] Regardless of the criteria that provenance tools are trying to meet, there is a blatant barrier to their effectiveness. These technology solutions are ultimately only as strong as their adoption, so they require intense outreach to put the tools in the hands of those documenting international crimes. Furthermore, without robust capacity-building programs that give those behind the camera the knowledge to capture what is relevant to a case, and those receiving the material the expertise to interpret the metadata, the potential of provenance technology will be crippled.[28]

Detection Technology

Just as deepfakes grow more sophisticated, so does the need for tools that can detect whether a piece of audiovisual content has been AI generated or manipulated. It is unlikely that the ICC will be able to develop, and keep up the maintenance of, their own state-of-the-art detection tools. For this reason, partnerships with leading companies can help fill this detection gap (the Office of the Prosecutor already has an ongoing collaboration with Microsoft). However, it is important to note three shortcomings that may likely arise.

Generally speaking, detection tools are still in their infancy and have yet to be deployed at scale in a sufficiently reliable fashion. While there has been dramatic progress in audio synthesis over the year 2023, audio biometrics, which could help identify whether someone uttered certain words, still lag behind. As Sam Gregory, executive director at WITNESS, puts it, "[i]t's no use having a tool to spot whether content is generated by one company when the same tool would give a false negative on fake audio created by one of the many other tools on the market."[29] When AI detection tools are trained on data not specific to an individual, their accuracy is similar to that of humans.[30] If trained on the biometrics of a person, such as President Zelenskyy,[31] their accuracy in determining the authenticity of a given audio can jump to 99%. This said, current deepfake detection technology has been largely optimized for facial recognition, and it is less adept at discerning tampering in nonhuman

elements of footage. This limitation is critical: the manipulation of structures or the environment in a video—like a doctored image of a bombed building or the insertion of a falsified object such as illegal ammunition in a civilian area—can significantly alter the course of an investigation.

Second, algorithms trained to spot inconsistencies in images, videos, or audios underperform in non-English languages, noisy environments, and complex scenarios like conflict settings. Grainy footage, low-resolution, or lacking a representative biometric sample will challenge computational assessments. Additionally, automated analysis is more effective when the file is as close as possible to the original version. Social media platforms, by compressing images and modifying file metadata upon uploading, can obscure the digital footprint, thereby confounding automated systems. As long as audiovisual content continues to be posted online, the digital footprint will get muddled, confounding automated systems. This is why detection algorithms must be paired with contextual analysis and corroborating material to ascertain the authenticity of content. To enhance the sensitivity of these models, there is a pressing need for datasets of manipulated videos and audio, specifically derived from social media. Collaboration with these platforms is then essential to help researchers refine detection algorithms. Similarly, the integration of biometric information is necessary. However, we should also understand that acquiring potentially sensitive data could put at risk those to whom we are trying to bring justice, should this information fall in the hands of malicious actors or oppressive governments. History teaches us that mass atrocities are often preceded by mass surveillance, enabled by the collection of data and the monitoring of targeted communities.

Last, from a legal standpoint, the deployment of automated detection technology raises complex questions. If such evidence is to be used in court, how will it be presented and understood by the parties? The fact that a model did not find manipulation does not mean a specific piece of content is not a deepfake; it just means that the system did not detect what it was looking for.[32] For example, it could be a deepfake generated in a new way that the detection tool had not yet been trained to identify. Legal practitioners will also have to grapple with the multiple variations of a manipulation. A video could be entirely synthetic, or the deception could lurk only in a fraction of the frames or the voice-over. Current models struggle with pinpointing targeted manipulations, such as the addition of an object or the alteration of a background. Furthermore, the credibility of a witness could be called into question as a sophisticated deepfake could shake the confidence in the authenticity of legitimate evidence. If it is possible to falsify one event, what is stopping any evidence from being doubted? This broad basis for mistrust threatens the core of international justice—the establishment of an objective truth.

Analytical and Forensic Expertise

Given these shortcomings, detection tools will most likely be better employed to conduct the initial triaging of audiovisual content, instead of aiming at producing conclusive verification reports. Ascertaining authenticity will require investing in human analysis in a twofold manner, by strengthening the knowledge and capacity of staff, particularly analysts and the judiciary, but also through secondment of expertise from governments, other accountability bodies, specialized institutions like Justice Rapid Response (JRR), and the private sector. Interpreting the events under question will also necessitate updating the Court Registry's roster to incorporate expert witnesses who are skilled in media synthesis and forensic analysis and are able to testify in court to explain provenance techniques such as digital signatures, provenance standards like C2PA, detection techniques, and other technical aspects related to the authenticity of evidence.

As deepfake technology becomes more accessible and convincing, the ICC and accountability institutions must pivot towards sophisticated approaches to provenance, detection, and expertise that can keep pace with the evolving technology. The sheer quantity of potentially AI-generated or manipulated content threatens to overwhelm investigators and analysts tasked with triaging audiovisual content, potentially leading to miscarriages of justice. We owe it to the victims and survivors of these crimes to invest now in preparing our international justice system for a future that is already here.

Raquel Vazquez Llorente *is the head of Law and Policy, Technology Threats and Opportunities at WITNESS. She is an international criminal lawyer specializing in how emerging technologies impact our trust in audiovisual content. She has over a decade of experience with digital evidence in conflict and crises.*

ENDNOTES

1. Megan O'Hearn, *But It Looks So Real! The Parallel Rise of Photography and Spiritualism*, JSTOR (Oct. 20, 2016), https://about.jstor.org/blog/but-it-looks-so-real-the-parallel-rise-of-photography-and-spiritualism/.

2. Ian J. Goodfellow et al., *Generative Adversarial Nets*, arXiv 1406.2661v1 [stat ML] (2014), https://arxiv.org/abs/1406.2661.

3. Sarah Cahlan, *How Misinformation Helped Spark an Attempted Coup in Gabon*, Wash. Post (Feb. 13, 2020), https://www.washingtonpost.com/politics/2020/02/13/how-sick-president-suspect-video-helped-sparked-an-attempted-coup-gabon/.

4. Sam Gregory, *The World Needs Deepfake Experts to Stem This Chaos*, WIRED (June 24, 2021), https://www.wired.com/story/opinion-the-world-needs-deepfake-experts-to-stem-this-chaos/; The Irrawaddy, *Myanmar Junta Accused of Using Deepfake Technology to Prove Graft Case Against Daw Aung San Suu Kyi*, Irrawaddy (Mar. 25, 2001), https://www.irrawaddy.com/news/burma/myanmar-junta-accused-using-deepfake-technology-prove-graft-case-daw-aung-san-suu-kyi.html.

5. Nathaniel Janowitz, *Venezuela Is Using Fake AI American Newscasters to Spread Disinformation*, Vice (Feb. 24, 2023), https://www.vice.com/en/article/z34jge/venezuela-ai-newscaster-disinformation; Adam Satariano & Paul Mozur, *The People Onscreen Are Fake.*

The Disinformation Is Real, N.Y. Times (Feb. 7, 2023), https://www.nytimes.com/2023/02/07/technology/artificial-intelligence-training-deepfake.html.

6. Rodrigo Soriano, *Las Imágenes de Xóchitl Gálvez Editadas con Inteligencia Artificial: Oportunidades y Peligros de Una Nueva Arma para la Comunicación Política*, El País (July 10, 2023), https://elpais.com/mexico/2023-07-10/las-imagenes-de-xochitl-galvez-editadas-con-inteligencia-artificial-oportunidades-y-peligros-de-una-nueva-arma-para-comunicacion-politica.html; Jorge Ramis, *Las Voces Generadas por IA se Vuelven Indistinguibles de las Humanas (También en Español)*," WIRED (Nov. 8, 2023), https://es.wired.com/articulos/voces-generadas-por-ia-indistinguibles-las-humanas-en-espanol.

7. Tom Simonite, *A Zelensky Deepfake Was Quickly Defeated. The Next One Might Not Be*, WIRED (Mar. 17, 2022), https://www.wired.com/story/zelensky-deepfake-facebook-twitter-playbook.

8. Jack Goodman & Mohanad Hashim, *AI: Voice Cloning Tech Emerges in Sudan Civil War*, BBC (Oct. 5, 2023), https://www.bbc.co.uk/news/world-africa-66987869.

9. Rania (@umyaznemo), *Do you see what I see? An Israeli refugee camp!*, X (Oct. 22, 2023, 8:42 PM), https://twitter.com/umyaznemo/status/1716173594818932808.

10. Aleksandra Wrona, *Are These Real Pics of a Massive Crowd of Israelis Demonstrating Support for Israel?*, Snopes (Oct. 23, 2023), https://www.snopes.com/fact-check/fake-photo-israel-crowds.

11. Shayan Sardarizadeh (@Shayan86), *This Video Certainly Doesn't Show a New Air Assault on Israel by Hamas Militants, Because It's Actually from the Video Game Arma 3*, X (Oct. 9, 2023, 1:44 PM), https://x.com/Shayan86/status/1711180889453953238?s=20.

12. Shayan Sardarizadeh (@Shayan86), *Some AI-Generated Images Are Being Shared in Relation to the Israel-Hamas Conflict*, X (Oct. 24, 2023, 3:54 PM), https://twitter.com/Shayan86/status/1716830625238544859/photo/1; Shahin Hazamy (@shahin_hazamy), *It's Genocide*, Instagram (Oct. 17, 2023), https://www.instagram.com/p/Cygt-uWLKwS/?hl=en.

13. *Preliminary Examination: Gabon, ICC-01/16*, ICC, https://www.icc-cpi.int/gabon.

14. *Information for Victims: Bangladesh/Myanmar*, ICC, https://www.icc-cpi.int/victims/bangladesh-myanmar.

15. *Venezuela: Investigating Crimes Against Humanity*, Clooney Found. for Just. (2023), https://cfj.org/the-docket/venezuela/#:~:text=Our%20team%20at%20The%20Docket,of%20security%20forces%20in%20Venezuela.

16. FIDH, *Mexico Requires the Support of the ICC to Eradicate Structural Impunity*, reliefweb (May 26, 2020), https://www.fidh.org/en/region/americas/mexico/mexico-requires-the-support-of-the-icc-to-eradicate-structural; FIDH, IDHEAS and Colectivo Solecito, "Hasta encontrarlos": enforced disappearances by security forces in Veracruz constitute crimes against humanity (Feb. 2022), https://www.fidh.org/IMG/pdf/enforced_disappearances_in_veracruz.pdf.

17. *Situation in Ukraine: ICC Judges Issue Arrest Warrants Against Vladimir Vladimirovich Putin and Maria Alekseyevna Lvova-Belova*, ICC (Mar. 17, 2023), https://www.icc-cpi.int/news/situation-ukraine-icc-judges-issue-arrest-warrants-against-vladimir-vladimirovich-putin-and#:~:text=Today%2C%2017%20March%202023%2C%20Pre,Ms%20Maria%20Alekseyevna%20Lvova%2DBelova.

18. *Darfur, Sudan: Situation in Darfur, Sudan, ICC-02//05*, ICC, https://www.icc-cpi.int/darfur.

19. Karim A. A. Khan KC, *Statement of ICC Prosecutor, Karim A. A. Khan KC, to the United Nations Security Council on the Situation in Darfur, Pursuant to Resolution 1593 (2005)*, ICC (July 13, 2023), https://www.icc-cpi.int/news/statement-icc-prosecutor-karim-khan-kc-united-nations-security-council-situation-darfur-0.

20. *Decision on the "Prosecution Request Pursuant to Article 19(3) for a Ruling on the Court's Territorial Jurisdiction in Palestine, ICC-01/18-143*, ICC (Feb. 5, 2021), https://www.icc-cpi.int/court-record/icc-01/18-143.

21. Karim A. A. Khan KC, *Statement of ICC Prosecutor on the Situation in the State of Palestine and Israel*, United Nations (Oct. 30, 2023), https://www.un.org/unispal/document/statement-of-icc-prosecutor-on-the-situation-in-the-state-of-palestine-and-israel/.

22. Beth Van Schaak, *Event: We Hold These Truths: How Verified Content Defends Democracies*, CSIS (Mar. 7, 2023), https://csis-website-prod.s3.amazonaws.com/s3fs-public/2023-03/230327_Verified_Content_Democracies.pdf?VersionId=2GcqVaJZzK6XqUaRoQNbD9ZQT6SRtfiI.

23. Lindsay Freeman & Raquel Vazquez Llorente, *How to Prepare the International Criminal Court for Our Digital Future*," OpinioJuris (Oct. 12, 2021), https://opiniojuris.org/2021/10/12/how-to-prepare-the-international-criminal-court-for-our-digital-future/.

24. Rome Statute of the International Criminal Court, art. 54.1.a, July 17, 1998, 2187 UNTS.

25. Lindsay Freeman & Raquel Vazquez Llorente, *Finding the Signal in the Noise: International Criminal Evidence and Procedure in the Digital Age*, 19 J. Int'l Crim. Just. 163 (2021).

26. Raquel Vazquez Llorente & Wendy Betts, *Coding Justice? The Tradeoffs of Using Technology for Documenting Crimes in Ukraine*," OpinioJuris (Oct. 22, 2023), https://opiniojuris.org/2022/10/22/coding-justice-the-tradeoffs-of-using-technology-for-documenting-crimes-in-ukraine/.

27. Chiara Gabriele, Kelly Matheson & Raquel Vazquez Llorente, *The Role of Mobile Technology in Documenting International Crimes: The Affaire Castro et Kizito in the Democratic Republic of Congo*, 19 J Int'l Crim. Just. 107 (2021).

28. Chiara Gabriele, Kelly Matheson & Raquel Vazquez Llorente, *Incorporating Digital Technology in the Investigation of International Crimes: Lessons from the Democratic Republic of Congo*," Just Sec. (Oct. 14, 2021), https://www.justsecurity.org/76780/incorporating-digital-technology-in-the-investigation-of-international-crimes-lessons-from-the-democratic-republic-of-congo.

29. Morgan Meaker, *Deepfake Audio Is a Political Nightmare*, WIRED (Oct. 9, 2023), https://www.wired.co.uk/article/keir-starmer-deepfake-audio.

30. Kimberly T. Mai et al., *Warning: Humans Cannot Reliably Detect Speech Deepfakes*, 18 PLoS One e0285333 (2023), https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0285333.

31. Matyáš Boháček & Hany Farid, *Protecting President Zelenskyy Against Deep Fakes*, arXiv (June 24, 2022), https://arxiv.org/pdf/2206.12043.pdf.

32. Mohamed Kambal (@Muhammed-Kambal), *Other Types of Forms of Manipulations Could Hatve Taken Place*, Twitter (Aug. 1, 2023, 7:59 PM), https://twitter.com/MuhammedKambal/status/1686451575978344470.