MAYER | BROWN

DEEPFAKES, SYNTHETIC DATA AND LEGAL DEFENSIBILITY:

The Next Frontier

TODAY'S PRESENTERS

PARTNER

MARCUS CHRISTIAN

PARTNER

AMBER THOMSON

CYBERSECURITY RESEARCHER, CENTER FOR LEGAL AND COURT TECHNOLOGY, WILLIAM & MARY LAW SCHOOL

DANIEL SHIN

AGENDA

1. Introduction
2. Deepfake Threats and Legal Exposure
3. Synthetic Data
4. Legal Defensibility
5. Practical Takeaways: Looking Ahead
6. Q&A

01

INTRODUCTION

WHY THIS TOPIC MATTERS

- Gen AI traffic experienced an explosive surge of over **890%** in 2024
  - In 2025, the average monthly number of Gen AI-related data security incidents **increased 2.5x,** now accounting for **14%** of all data security incidents across SaaS traffic
  - Synthetic content is increasingly used in social engineering attacks
- Rapid growth of Generative AI tools that make it easy to manipulate and create hyper-realistic images, videos and voices
- Rise in deepfake-enabled fraud and deception
  - Weaponized to enable cyber bullying, harassment, and damage reputations
  - Disproportionate effect on young people and women
- Some legal and regulatory momentum over the past year to address deceptive uses of AI and synthetic media
- Urgent need for awareness and safeguards

SETTING THE SCENE

ARUP

British engineering group Arup lost approximately $25 million after scammers used AI-manipulated “deepfakes” to falsely pose as the group’s CFO and request transfers from an employee to bank accounts in Hong Kong.

For almost a whole day, AI generated pornographic images of Taylor Swift circulated social media platforms.

Within hours, some images were seen more than 45 million times and accrued thousands of shares and likes before eventually being taken down.

02

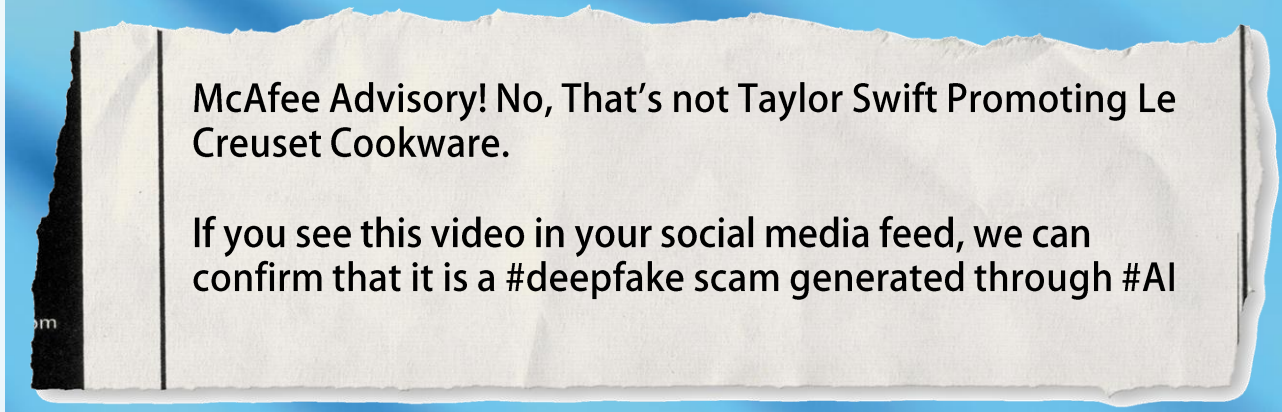
DEEPFAKE THREATS AND LEGAL EXPOSURE

WAR STORIES

- Business email compromise 2.0
  - Deepfake videos and voice impersonations
  - Phishing
  - Blackmail
  - Misinformation campaigns
  - Synthetic identity fraud
  - Social engineering amplification
- Generative AI could enable fraud losses to reach **$40 billion** in the U.S. by 2027.*

CEO of world's biggest ad firm targeted by deepfake scam

Exclusive: fraudsters impersonated WPP's CEO using a fake WhatsApp account, a voice clone and YouTube footage used in a virtual meet



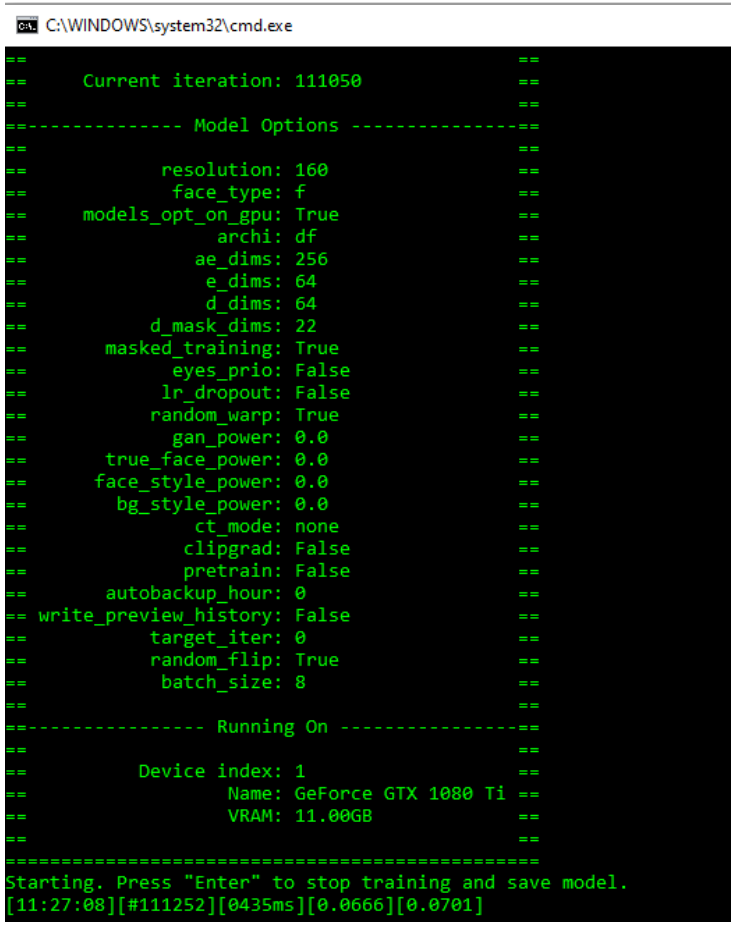
McAfee Advisory! No, That's not Taylor Swift Promoting Le Creuset Cookware.

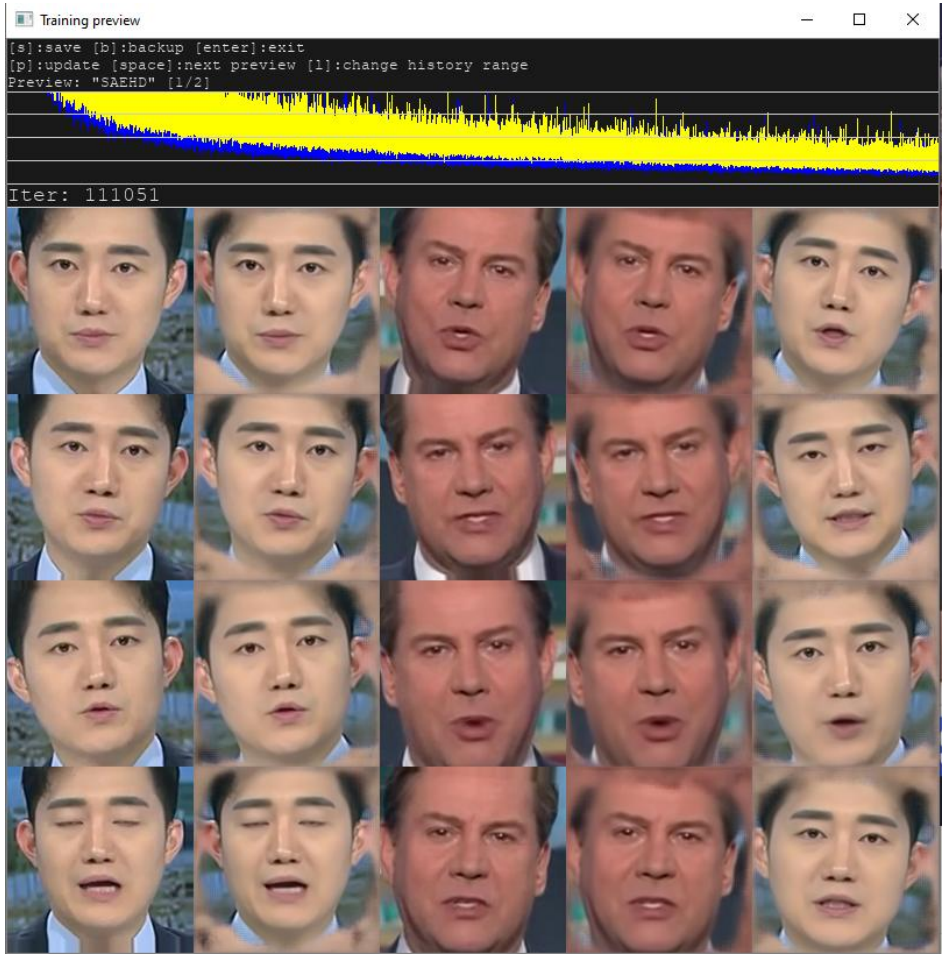
If you see this video in your social media feed, we can confirm that it is a #deepfake scam generated through #AI

* https://www.deloitte.com/us/en/insights/industry/financial-services/deepfake-banking-fraud-risk-on-the-rise.html

FACE SWAP DEEPFAKES

FACE SWAP DEEPFAKES





FACE SWAP DEEPFAKES

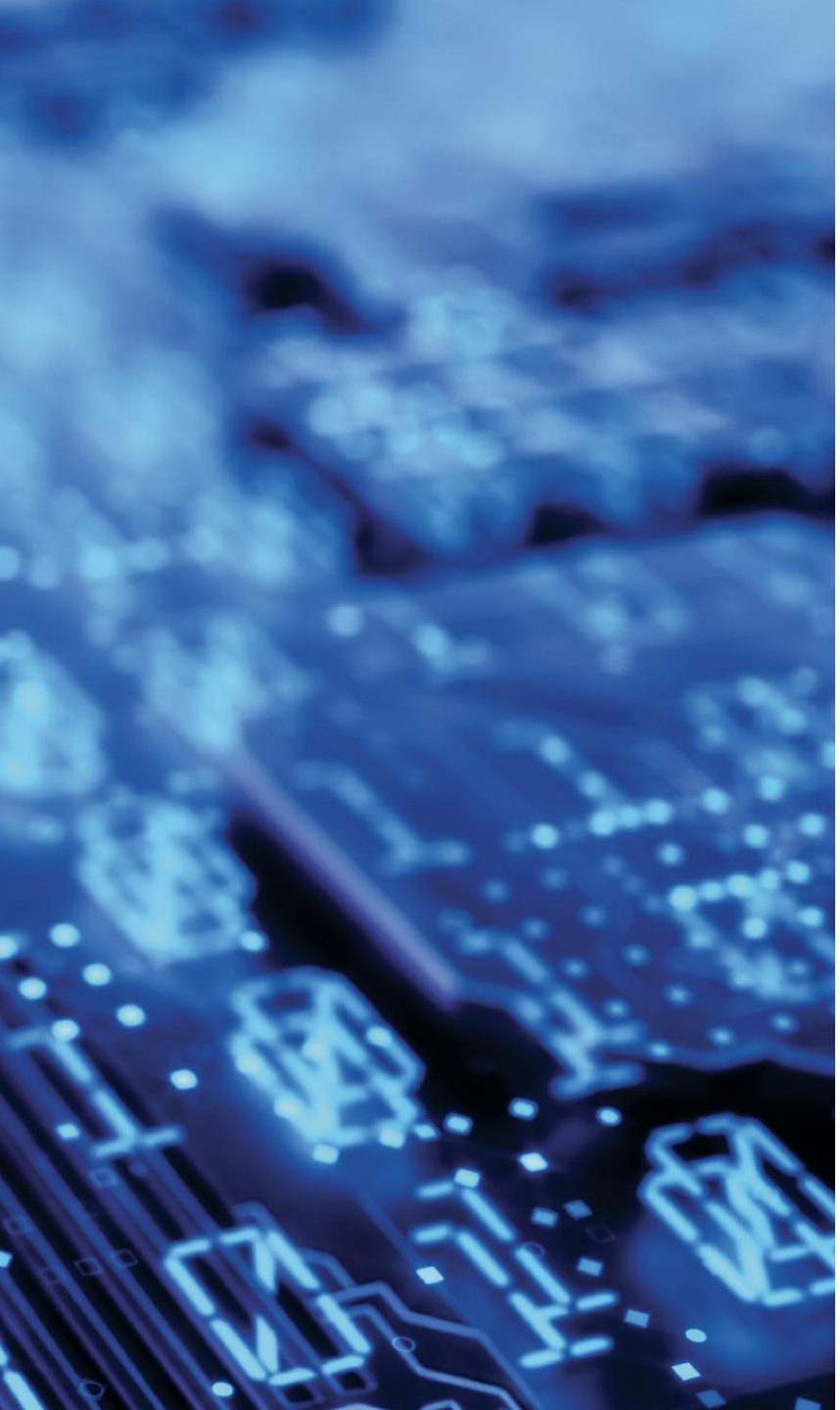
TEXT-TO-VIDEO DEEPFAKES

2024 HURRICANE HELENE

2024 HURRICANE HELENE

2024 HURRICANE HELENE

Image Forensics by Professor Hany Farid, University of California at Berkeley.

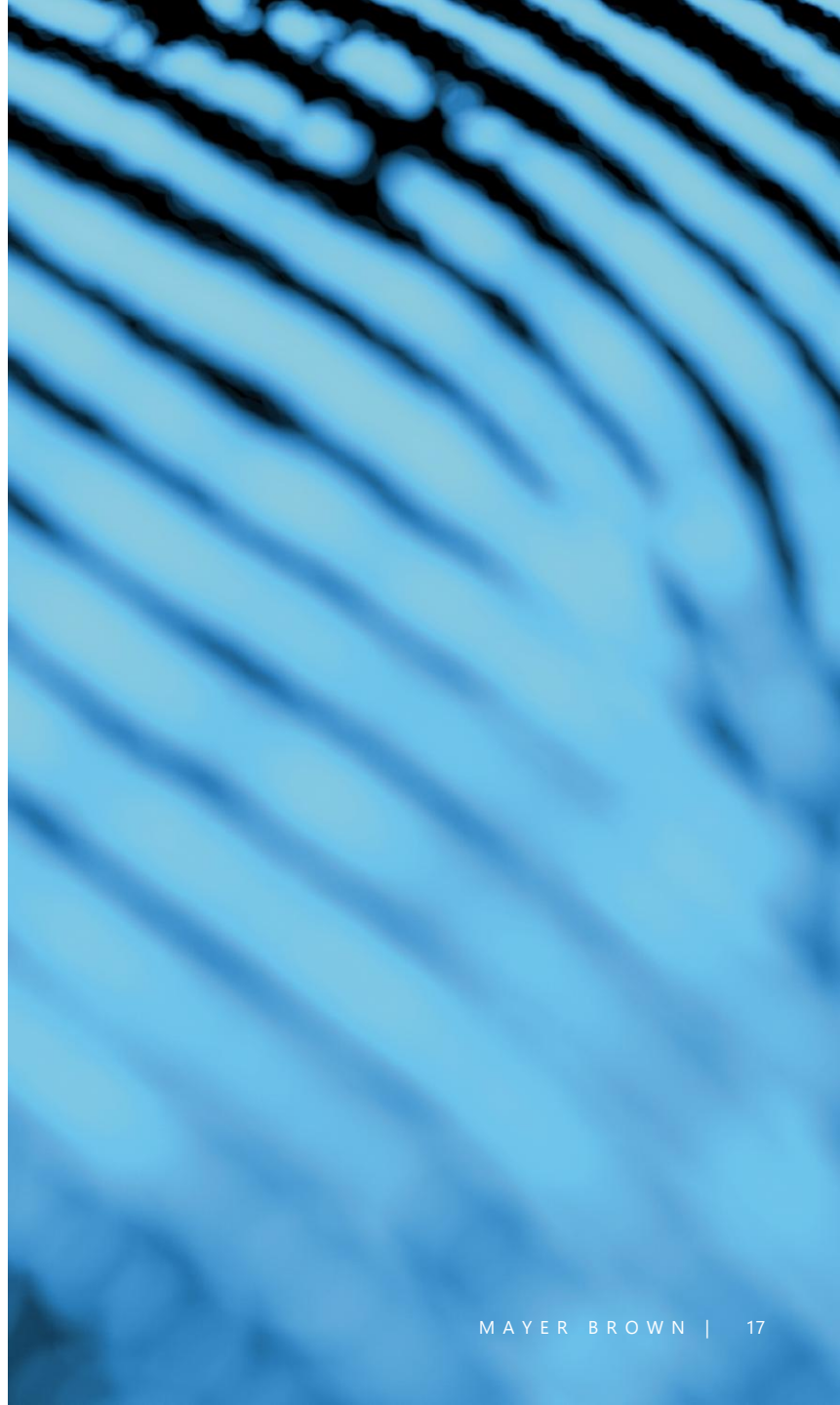


LEGAL RISKS AND EXPOSURE

- Core risks:
  - Defamation
  - Tort liability
  - IP violations
  - Cybersecurity / data breach implications
- Additional exposure:
  - Right of publicity
  - Biometric privacy laws (e.g., BIPA, NY §50-f)
  - IP infringement using copyrighted faces, trademarks, or voices

DETECTION CHALLENGES

- Detection remains difficult due to rapid model evolution.
- Current detection and watermarking methods remain constrained by technical and practice limitations.
    - Watermarking is useful but not sufficient on its own. It should be part of a layered approach.
- Although most major U.S.-based generative AI service employ watermarking-based content labeling to allow future data provenance analysis not all generative AI services employ watermarking
    - E.g., China-based text-to-image or text-to-video services.
    - Open-source generative AI tools largely do not employ watermarking techniques to synthesized media.
    - Watermarking-based content labeling must be done to both synthesized and “real” media in the medium term.

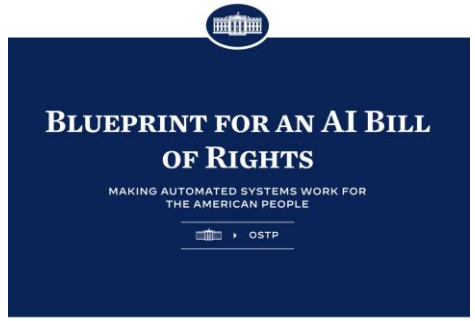


REGULATORY LANDSCAPE

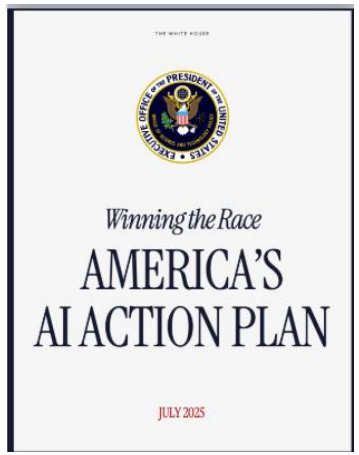
- Currently, there is no **comprehensive federal legislation** that regulates the development of AI or specifically prohibits or restricts its use.
- Legislative momentum:
  - FTC Impersonation Rule
  - FTC crack down on deceptive AI claims and scheme
  - FCC ban on AI-generated robocalls
  - State laws criminalizing deceptive media (*e.g.*, TX SB 751, CA AB 730, NJ 2025 law)
  - Federal legislation (*e.g.*, AI Training Act, National AI Initiative Act)
  - Colorado AI Act
  - EU AI Act
  - TAKE IT DOWN Act aims to combat non-consensual synthetic sexual imagery

NON-BINDING FRAMEWORKS AND GOVERNMENT INITIATIVES

- AI Bill of Rights
- AI Action Plan
  - Accelerating innovation, building AI infrastructure, and leading in international diplomacy and security
- NIST AI Risk Management Framework
- The Presidential AI Challenge
- The White House Task Force on Artificial Intelligence Education



Guidebook for Participation



EXECUTIVE ORDERS

- Promoting the Export of the American AI Technology Stack | 7/23/2025
- Accelerating Federal Permitting of Data Center Infrastructure | 7/23/2025
- Preventing Woke AI in the Federal Government | 7/23/2025
- Advancing Artificial Intelligence Education for American Youth | 4/23/2025
- Restoring Common Sense to Federal Procurement | 4/15/2025
- Removing Barriers to American Leadership in Artificial Intelligence | 1/23/2025
- Initial Rescissions of Harmful Executive Orders and Actions | 1/20/2025
- Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government | 12/3/2020
- Maintaining American Leadership in Artificial Intelligence | 2/11/2019

03

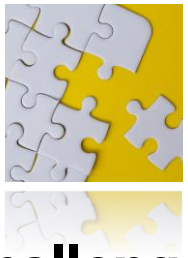
SYNTHETIC DATA

SYNTHETIC DATA

What is it?

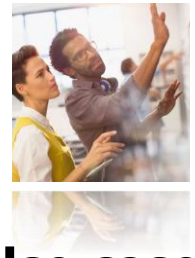
Synthetic data is **artificial data** designed to mimic real-world data. It's generated through statistical methods or by using AI techniques like deep learning and Gen AI.

Can also act as a **placeholder** for test data, especially where real data is limited, sensitive, or risky to use.



Challenges

Bias, model collapse, trade-off between accuracy and privacy, verification

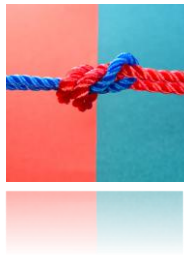


Use cases

Finance – Simulate customer transaction data to test fraud detection systems while protecting actual account details.

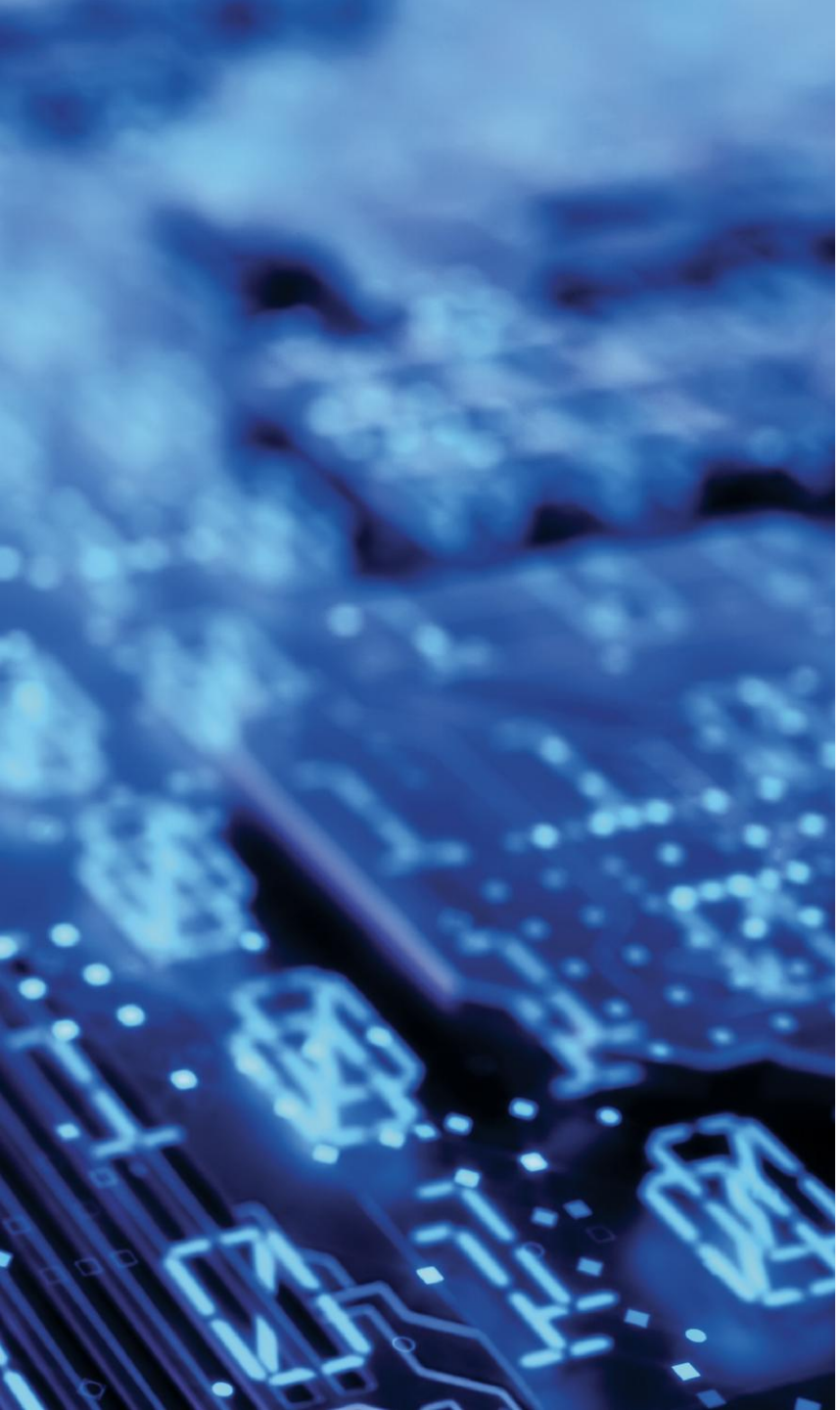
Healthcare - Generate patient-like data for research and model training without exposing real patient records.

Retail – Create synthetic shopper profiles to optimize store layouts.



Dual-use context

Legitimate privacy-preserving training versus abuse cases and the regulatory gray zone when synthetic becomes deceptive.



CASE STUDIES

- "*The creation or distribution of synthetic content is not inherently illegal; however, synthetic content can be used to facilitate crimes, such as fraud and extortion.*" – FBI IC3
- Synthetic identities and records
  - Attackers create entirely or partially fabricated people/accounts and use them to open accounts, approve transactions, or inflate metrics.
- Synthetic data sets
  - Fake data sets can be created to support due diligence, audits, or internal controls – intended to appear "audit proof".
- In 2023 Charlie Javice was arrested on charges that she defrauded JPMorgan when she sold her company.
  - Months after the sale, JPMorgan learned that the actual number of people with accounts was around 300,000 and not 4 million users.
  - A mathematics professors was paid to create a synthetic data file of four million fake people.

RISK & COMPLIANCE CHALLENGES

Risks

- Sloppy data in → Sloppy data out
- Model collapse / drift
- Abuse scenarios

Compliance Challenges

- Governance blind spots
- Thresholds for when synthetic data triggers consumer protection or sector-specific obligations.

REGULATORY SCRUTINY

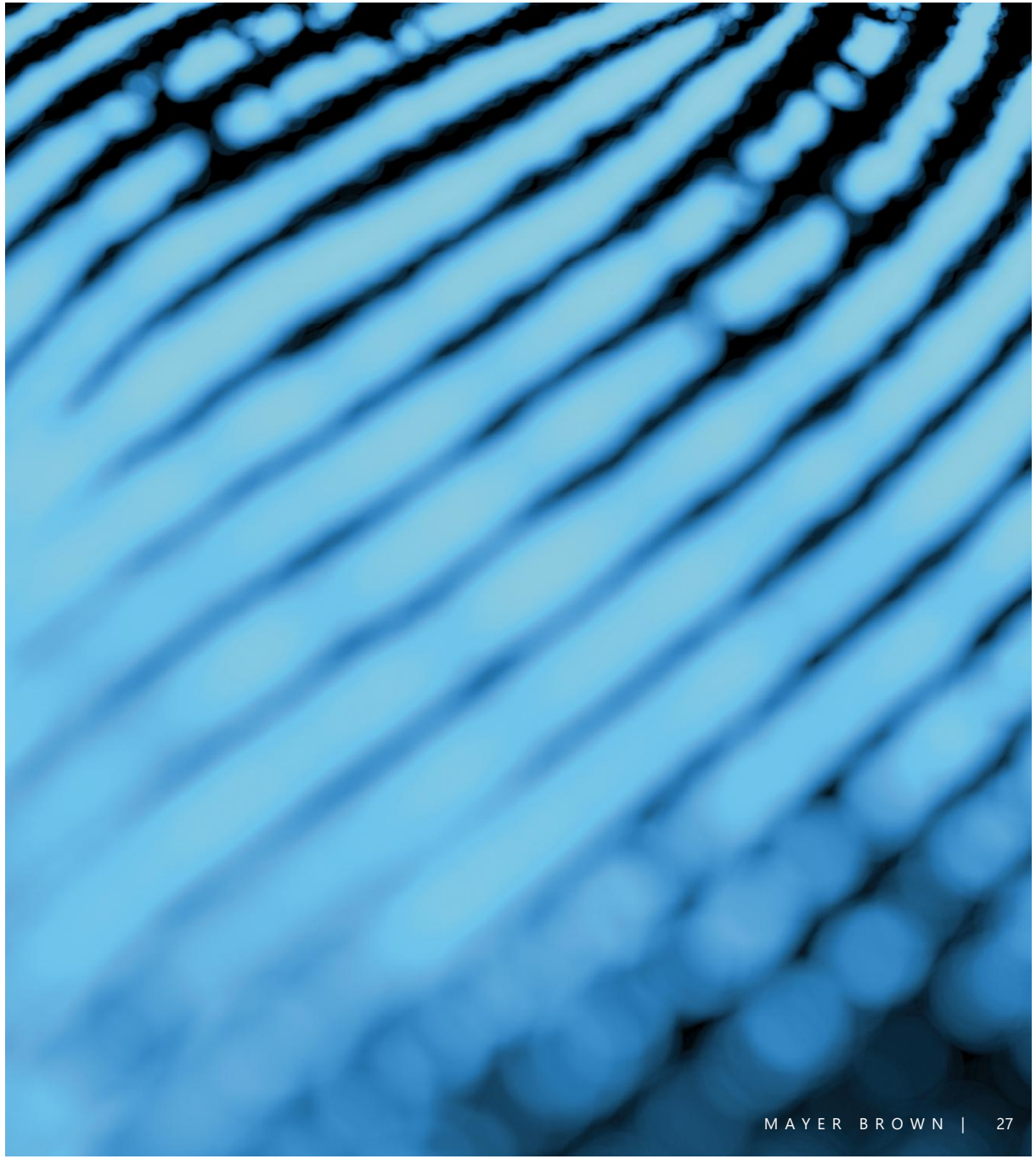
- Fragmentation of U.S. privacy laws can make it challenging to determine the legal status of synthetic data in various contexts.
- Synthetic data is getting increasing attention from regulators globally. But why?
  - Privacy mitigation (reducing reliance on real personal data)
  - Questions about consent, data provenance, disclosure, transparency
  - Use of synthetic data to train high-risk AI models has raised concerns
  - Push for clearer disclosures when synthetic data used in decision-making context

04

LEGAL DEFENSIBILITY

LITIGATION RISKS

- Authentication
  - FRE 901 and 902(13)/(14)
- Evidence integrity
  - Chain-of-custody complications with manipulated evidence.
- Discovery challenges
  - Expert testimony and Daubert challenges to AI detectors.
- **Example**: “deepfake alibi” defenses in criminal and civil matters



BOARD & C-SUITE RESPONSIBILITY AND OVERSIGHT

- Ensure proper governance mechanism is employed to both commercial and open source-based AI systems.
- Counsel's role in aligning AI governance with ISO/IEC 42001 and NIST AI RMF.
  - Interpret and operationalize legal and regulatory requirements.
  - Advise on risk management and accountability structures.
  - Monitor evolving standards.
- Integrating deepfake and synthetic data risks into ERM.
  - Embed emerging synthetic media threats into enterprise risk frameworks.
  - Strengthen governance, detection, and response capabilities

KEY QUESTIONS DIRECTORS SHOULD ASK ABOUT AI USE AND DETECTION CONTROLS

- **Strategic Alignment**
  - How is AI being used to support our business strategy and operational goals?
- **Governance and Oversight**
  - Do we have a formal AI governance framework in place?
  - Who is accountable for AI risk management?
  - Are board members receiving regular updates on AI-related risks and opportunities?
- **Risk and Compliance**
  - Have we assessed legal, ethical, and reputational risks associated with our AI systems?
  - Are we complying with emerging AI regulations?
  - Do we have protocols for managing bias, fairness, and explainability in AI outputs?
- **Data Integrity and Privacy**
  - What types of data are used to train our AI models?
  - Are synthetic data or deepfakes used in any part of our operations or testing?
  - How do we ensure compliance with data protection laws?
- **Detection and Monitoring Controls**
  - What controls are in place to detect misuse, manipulation, or unintended behavior in AI systems?
  - Do we use tools to identify deepfakes or synthetic media threats?
  - How do we monitor AI performance over time?

KEY QUESTIONS DIRECTORS SHOULD ASK ABOUT AI USE AND DETECTION CONTROLS (CONT.)

- **Incident Response and Accountability**
  - Do we have an incident response plan for AI-related failures or misuse?
  - How would we respond to a deepfake attack or synthetic fraud targeting our organization?
  - Are there escalation paths and communication protocols for AI-related incidents?

- **Third-Party and Vendor Risk**
  - Are we evaluating AI risks in our vendor ecosystem?
  - Do third-party tools or platforms we use deploy AI, and are they subject to our governance standards?

- **Transparency and Stakeholder Trust**
  - Are we transparent with customers, employees, and regulators about how we use AI?
  - Do we have clear disclosures and opt-out mechanisms for AI-driven decisions?

- **Board Readiness**
  - Do directors have the training and resources to understand AI risks and opportunities?
  - Should we consider adding AI expertise to the board or advisory committees?

BUILDING DEFENSIBLE FRAMEWORKS

Develop internal policies for generative content use and monitoring.

Create incident response playbooks addressing AI-driven deception.

Maintain documentation and audit trails as part of defensible AI governance.

Ensure readiness for regulatory inquiries and courtroom scrutiny.

CASE STUDIES / LESSONS LEARNED

- Corporate fraud: Arup deepfake video call revealing internal control failure.
- Reputational harm: Celebrity deepfake images and platform response.
  - Some Korean schools don’t publish school photo book to prevent deepfake threats against students.
- Evidence manipulation: synthetic voice recordings surfacing in discovery.

In these cases, what was missing?
Policy, detection, or legal readiness?

05

PRACTICAL TAKEAWAYS: LOOKING AHEAD

TAKEAWAYS

- Adopt provenance and authenticity verification technologies early.
- Maintain multidisciplinary response teams across Legal, InfoSec, and Comms.
- Track evolving laws and FTC/FCC/EDPB enforcement trends.
- Prepare for discovery and compliancy scrutiny involving AI-generated content.
- Employ data minimization, especially for personal data that could be used to create deepfake media (photos, audio, and video media).

06

Q&A

MAYER | BROWN