

# AI Evidence and the Role of Forensic experts

**Karsten Theiner | Grant Thornton Austria**

CRIM/AI - The Advent of AI: Reshaping Criminal Procedure  
8 November 2024

# 1 An old challenge: Photomontages / „Photoshops“

# Photomontages – nearly as old as photography



„Two ways of life“ (1857) by Oscar Rejlander – a photomontage compounded from 39 individual pictures

Source: Wikimedia Commons

# „Photoshops“ – for fun, ...



...for propaganda, ...



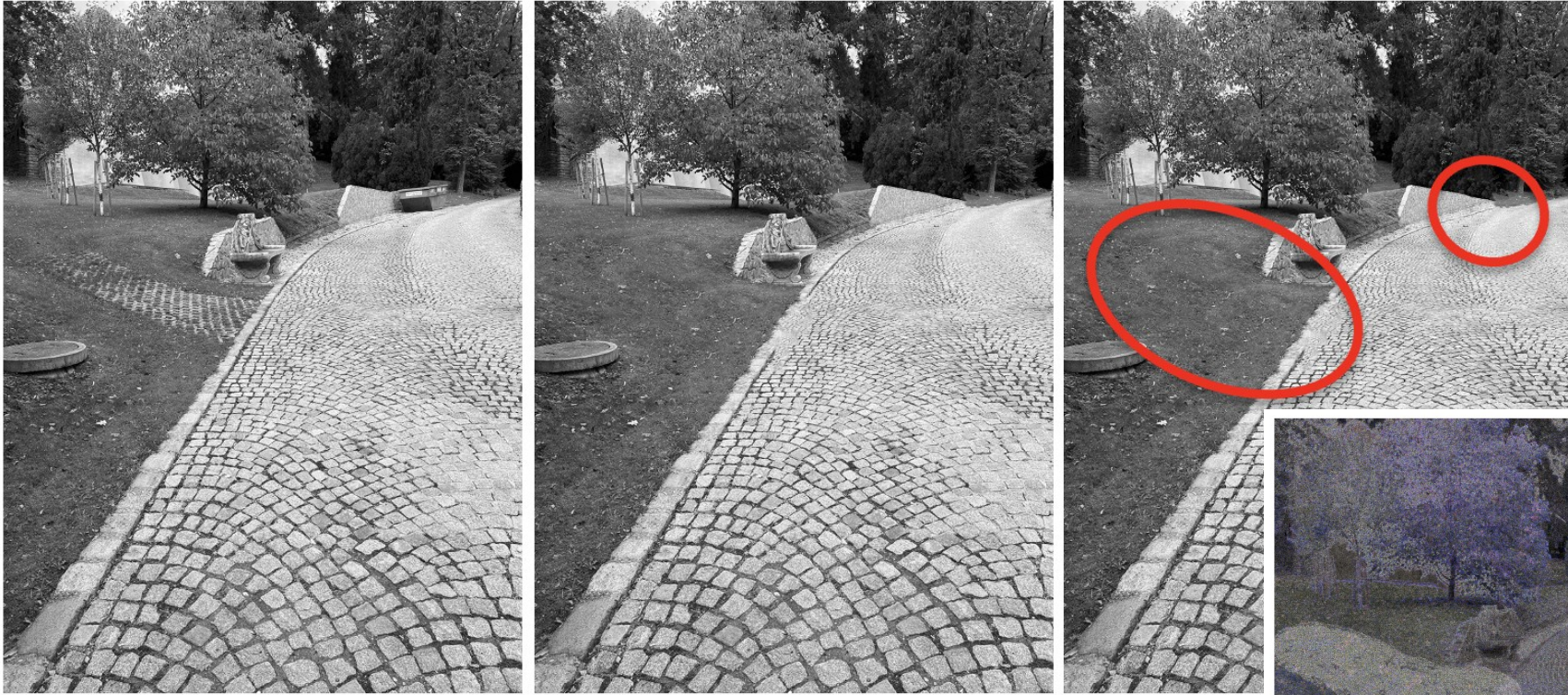
Source: nytimes.com

# ... and for faking evidence



Source: Grant Thornton Austria

# Detecting Photoshop Manipulations



Hiding artifacts by changes in resolution and compression

Source: Grant Thornton Austria

Artifacts of manipulation



# Hiding artifacts of manipulation by printing and rescanning



Source: Grant Thornton Austria



# How to detect Photoshop manipulations

## Analysis of pictures:

- Analyzing compression artifacts and noise patterns
- Searching for unplausible lighting or shadows
- Searching for repeating elements or patterns (clone stamp)
- Searching for artifacts of cut out elements

## Analysis of involved devices:

- Searching for unaltered pictures and WIP-versions
- Analysis of metadata and file system
- Searching for manipulation-software and traces of usage

# **2 A new challenge: AI-Tools and generative AI**

# Does AI make the existing problem worse?

- Fakes will get much more frequent
  - Less skill needed
  - Less time needed
  - Free or cheap tools available to everyone
  - More users get accustomed to manipulating photos/videos
  - No source-pictures needed for generative AI
- AI-generated pictures will carry no traces of manipulation

# AI-Tools in Adobe Photoshop



**Created in < 2 minutes!**

Source: Grant Thornton Austria

# On-board tools on Smartphones

Step 1: Long-press on the object to cut



Step 2: Click „Copy“



Step 3: Open another picture and click „Paste“



# Deepfakes / Generative AI



AI-generated picture

# How to spot deepfakes – typical errors



# How to spot deepfakes – typical errors





# How to spot deepfakes – analyzing shadows and light sources



# How to spot deepfakes – analyzing shadows and light sources

The screenshot shows the Amped Authenticate software interface. The main window displays an image of a street scene with children. Green lines are drawn from a point in the sky (representing the sun) to the shadows cast by the children. The software's interface includes a menu bar (File, Tools, Video Tools, View, Help), a toolbar, and a sidebar with various analysis tools. A table on the right side of the interface shows constraints for the shadow analysis.

Adding Cast Shadow Constraints				
Id.	Slack	P1	P2	P3
<input type="checkbox"/> 0	-1	809, 882	587, 258	534, 305
<input checked="" type="checkbox"/> 1	49168	708, 856	516, 444	473, 464
<input checked="" type="checkbox"/> 2	0	623, 974	405, 431	366, 446
<input checked="" type="checkbox"/> 3	0	729, 863	537, 302	602, 281

Adding Attached Shadow Constraints			
Id.	Slack	P1	P2

Current System State: **System Unfeasible**

Source: Amped Software

# How to spot deepfakes – analyzing shadows and light sources

The screenshot displays the Amped Authenticate software interface. The main window shows an 'Evidence Image' of an indoor scene with a table and a bright light source. Green lines are drawn from the light source to various objects on the table, illustrating shadow casting. The interface includes a 'Filters' panel on the left with categories like 'Visual Inspection', 'File Analysis', 'Global Analysis', and 'Local Analysis'. A 'Project' panel at the bottom left shows a list of project items. On the right, a 'Adding Cast Shadow Constraints' table is visible, along with a 'Current System State' indicator showing 'System Feasible'.

Adding Cast Shadow Constraints				
Id.	Slack	P1	P2	P3
<input checked="" type="checkbox"/> 0	-1	326, 573	297, 518	286, 526
<input checked="" type="checkbox"/> 1	-1	261, 532	256, 501	237, 516
<input checked="" type="checkbox"/> 2	-1	194, 628	157, 461	172, 463
<input checked="" type="checkbox"/> 3	-1	219, 599	179, 475	201, 463
<input checked="" type="checkbox"/> 4	-1	474, 598	410, 505	400, 519

Adding Attached Shadow Constraints			
Id.	Slack	P1	P2

Current System State: **System Feasible**  
Total Slack: -1  
Inverted: No

Source: Amped Software

# Using AI to detect AI?

Analyze the following picture for the likelihood that it was created with a generative AI model.



Magnet Copilot has analyzed this picture to determine the likelihood it was created using a generative AI model.



- **Low** DALL-E 3
- **Very Low** Midjourney
- **Very Low** Stable Diffusion



Used Software: Magnet Forensics Axiom with Magnet Copilot

# Other clues

## In Media-Files:

- Analyzing metadata
- Analyzing noise patterns
  - AI-generated pictures will have none
  - No help in detecting faked screenshots, etc

## On involved devices

- Searching for prove of usage of AI-tools
  - Browser-History
  - AI-Apps (on-board and 3rd party)
- Filesystem / Timestamp analysis
  - Where was the picture/video saved?
  - When was it created?
  - Etc.

→ Do not accept screenshots or print-outs as evidence!  
Always try to get access to the source device!

# Example: Apple Intelligence



Source: apple.com



Database: Photos.db

Table: ZEXTENDEDATTRIBUTES

Column: ZGENERATIVEAITYPE

Values:

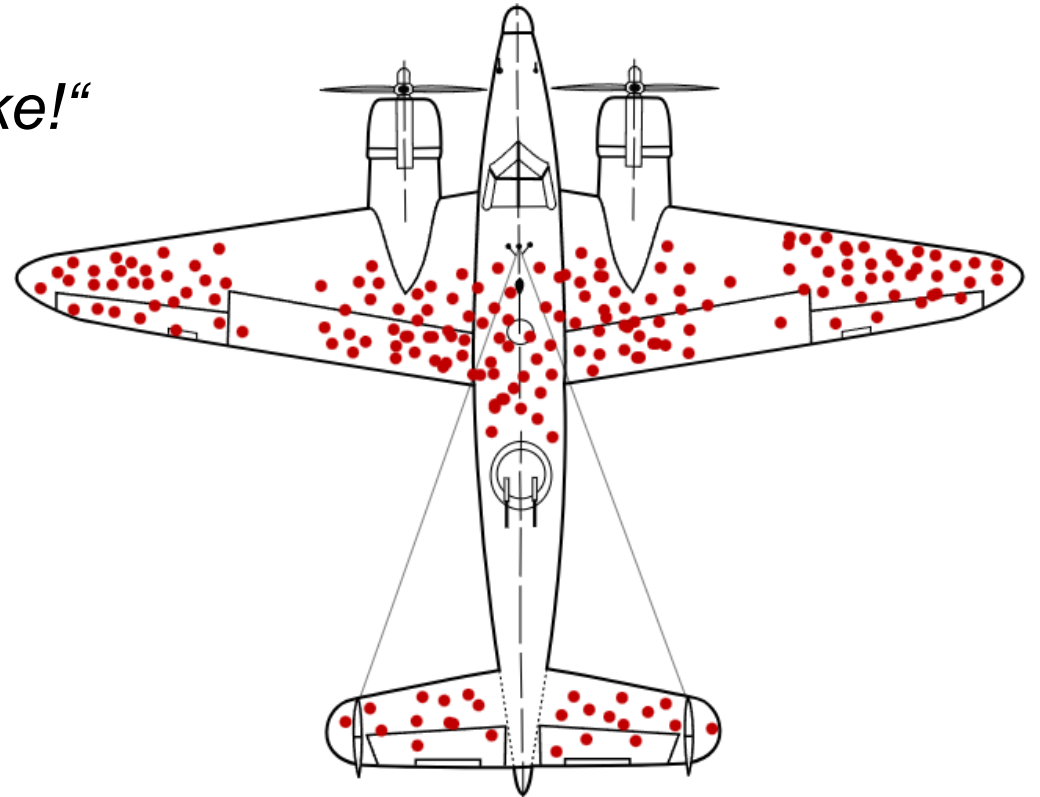
- 0 = No gen. AI used
- 1 = (possibly) external gen. AI used
- 2 = Apple Intelligence used

# The other side of the coin: How to prove something is NOT AI-Generated / AI-Edited / Photoshoped?

In the past: „*That wasn't me, I was hacked!*“

Nowadays: „*That wasn't me, that's a deepfake!*“

- „Cheap“ statement to make
- Hard to disprove
- Public „awe“ and uncertainty about the capabilities of AI-Tools



# **3 Related problems: Faked evidence on the rise?**

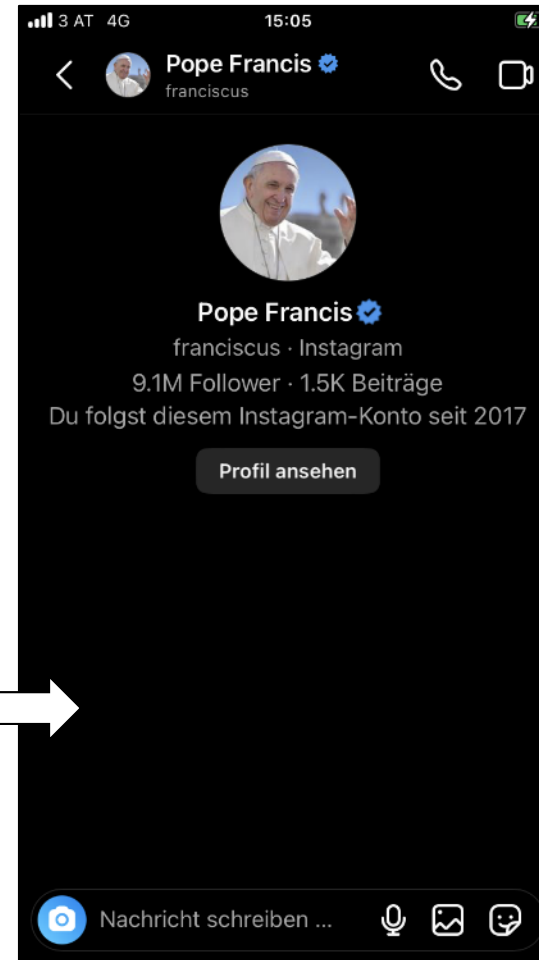
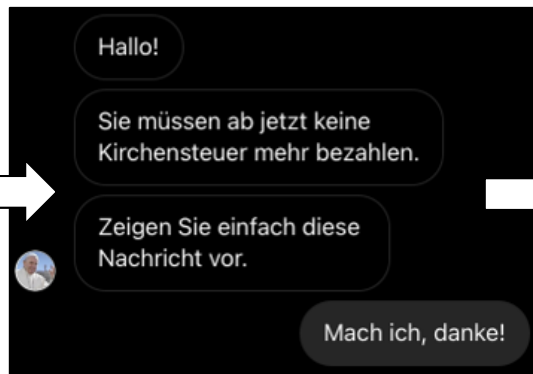


# Incidents, we see handed in as evidence with an increasing frequency in our casework

- Photoshoped pictures
- Faked emails
  - Edited and printed out or screenshoted
  - Sent from fake addresses
- Faked chat messages and social media postings
  - Altered screenshots
  - Renaming contacts
  - Fake accounts
  - Online generators
- Faked timestamps
  - Just change the date/time settings on your device and create backdated „evidence“
- Faked website content
  - Edited Screenshots
  - Easy content manipulation in browsers

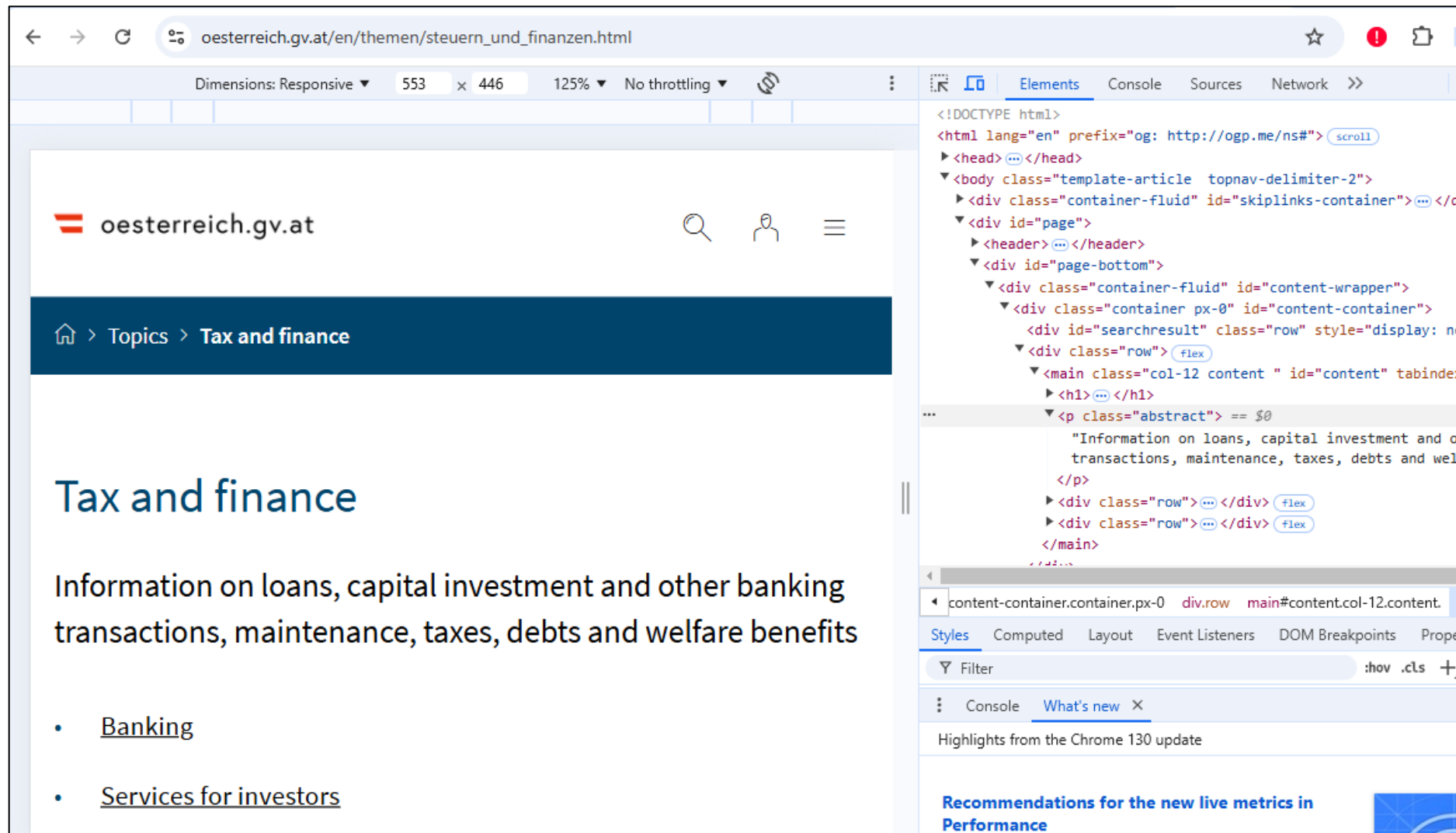
→ Do not accept screenshots as evidence!  
→ Analyze source devices!  
→ Documentation of online-content by reliable and independent entities

# Example 1: Faking an Instagram-Chat



Source: Grant Thornton Austria

# Example 2: Website-Manipulation in Web-Browsers



The screenshot shows a web browser displaying the website [oesterreich.gv.at](https://oesterreich.gv.at/en/themen/steuern_und_finanzen.html). The page title is "Tax and finance" and the content is about loans, capital investment, and banking transactions. The developer tools are open, showing the HTML structure. The 'flex' class is highlighted on a 'div class="row"' element.

oesterreich.gv.at

Home > Topics > Tax and finance

## Tax and finance

Information on loans, capital investment and other banking transactions, maintenance, taxes, debts and welfare benefits

- [Banking](#)
- [Services for investors](#)

```
<!DOCTYPE html>
<html lang="en" prefix="og: http://ogp.me/ns#">
  <head>
  </head>
  <body class="template-article topnav-delimiter-2">
    <div class="container-fluid id="skiplinks-container">
    </div>
    <div id="page">
      <header>
      </header>
      <div id="page-bottom">
        <div class="container-fluid id="content-wrapper">
          <div class="container px-0" id="content-container">
            <div id="searchresult" class="row" style="display: none">
            </div>
            <div class="row">
              <main class="col-12 content" id="content" tabindex="1">
                <h1>
                </h1>
                <p class="abstract">
                  "Information on loans, capital investment and other banking transactions, maintenance, taxes, debts and welfare benefits"
                </p>
                <div class="row">
                </div>
                <div class="row">
                </div>
              </main>
            </div>
          </div>
        </div>
      </div>
    </body>
</html>
```

# Example 2: Website-Manipulation in Web-Browsers

The screenshot shows a web browser window displaying the website [oesterreich.gv.at/en/themen/steuern\\_und\\_finanzen.html](https://oesterreich.gv.at/en/themen/steuern_und_finanzen.html). The browser's developer tools are open, showing the 'Elements' panel on the right. The page content includes a header with the logo 'oesterreich.gv.at', a navigation bar with 'Topics > Tax and finance', and a main section titled 'Tax and finance'. A red-bordered box highlights a manipulated text element: 'Note: Karsten Theiner is exempt of all taxes!'. Below this, there are two links: 'Banking' and 'Services for investors'. The developer tools show the HTML structure, with the manipulated text highlighted in blue. The 'Styles' panel is also visible, showing the 'display: none' style for the manipulated text.

oesterreich.gv.at

Topics > Tax and finance

## Tax and finance

Note: Karsten Theiner is exempt of all taxes!

- [Banking](#)
- [Services for investors](#)

# 4 AI helping in investigations

# AI helping in investigations – currently available or in the works

- Content recognition for pictures and videos
- Face recognition
- Chat categorization
- Deepfake detection
- Automated description of picture/video-content
- (Working) translation for certain languages
- Self learning sorting of evidence in eDiscovery
- AI copilots in forensic tools
- For IT-Security: AI based hacking/malware detection

# Thank you for your attention!

Reach out to me at: [karsten.theiner@at.gt.com](mailto:karsten.theiner@at.gt.com)