



UNIVERSITÄT
DES
SAARLANDES

INSTITUT FÜR
RECHTSINFORMATIK

Deepfakes – A Use Case for the Certification of AI Systems?

CARAT Project | 28. April 2026 | São Paulo

Prof. Dr Georg Borges



- Chair of Civil Law, Legal Informatics, German and International Business Law and Legal Theory, Saarland University
- Director of the Institute of Legal Informatics, Saarland University
- Judge, Higher Regional Court of Hamm (2012–2015)
- Member of the Hörst-Görtz Institute for IT Security (HGI) (2005–2015)
- Member of the Board, EDV-Gerichtstag e.V. [German Association for eJustice]
- Member of the Board, Stiftung Datenschutz [Data Protection Foundation]
- Member, EU Commission Expert Group on “Liability and new technologies, New technologies formation” (2018–2020)
- Member, EU Commission “Expert Group on B2B Data Sharing” (2022–2025)
- Distinguished Visiting Professor, University of Johannesburg (since 2023)
- Visiting Professor, Keio University, Tokyo (since 2024)



Jonas Herrmann, Dipl.-Jur.



- Member of Prof. Dr Georg Borges' research group since 2023
- First State Examination 2025
- PhD student under Prof. Dr Georg Borges since 2026



Agenda

I. Introduction

1. (AI) Deepfakes
2. A case study of deepfakes

II. Deepfakes in the AI Act

1. Definition in Art. 3 Nr. 60 AI Act
2. Transparency obligations under Art. 50 AI Act
3. Addition of a prohibition in Art. 5 AI Act

III. Certification

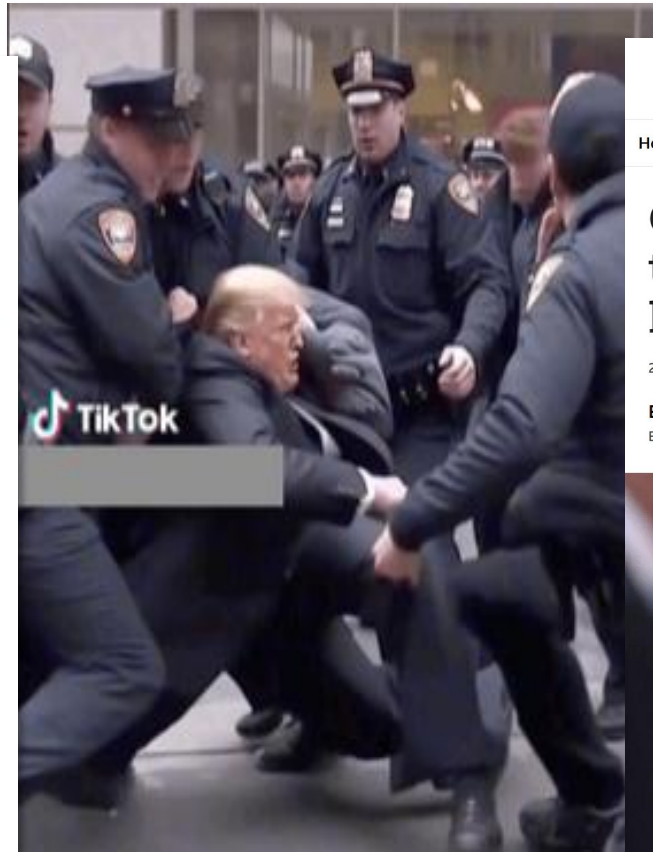




I. Introduction

1. (AI) Deepfakes

- Any of various media, *esp.* a video, that has been digitally manipulated to convincingly replace one person's likeness with that of another, often used maliciously to show someone doing something that he or she did not do (*Oxford English Dictionary*)
- "Deep-fake": A portmanteau of 'deep learning' and 'fake'



BBC

Home News Sport Business Technology Health Culture Arts Travel Earth Audio Video Live

German outcry over deep fake porn targeting actress prompts bid to change law

24 March 2026

Share Save Add as preferred on Google

Bethany Bell
Berlin reporter



2. A case study on deepfakes

Student A uses the AI system “FaceForge AI”, developed by **provider P**, which is operated and made available via the platform of **deployer D**. **A** uploads publicly available photos and lecture clips of **Professor X** and, just for fun, creates a video showing **X** in a lecture saying disparagingly that the students are “too stupid for law anyway” and that he awards poor marks “on a whim”.

A sends the video to **Student B**. A few weeks later, **B** is annoyed because **X** has given him a poor mark for his term paper and publishes the deepfake in a student chat group and on a social media platform. This triggers a major online backlash.

Relevant areas of law

General right of personality (APR)	Data protection law	Regulation under the AI Act
Copyright	Criminal law	Media and press law





II. Deepfakes in the AI Act

1. Definition

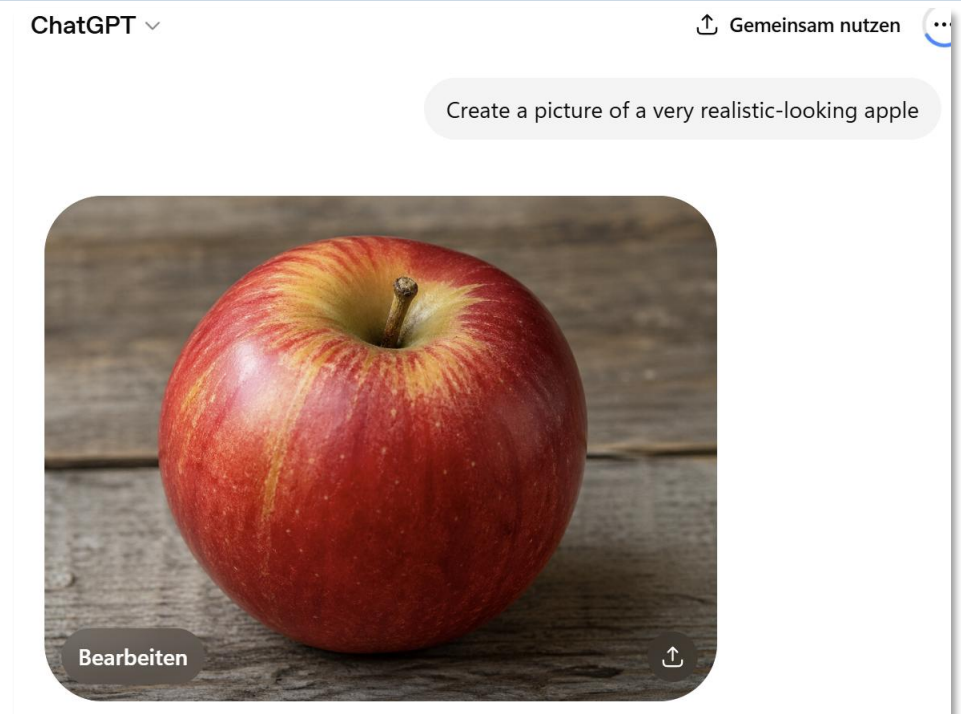
Art. 3 AI Act *Definitions*

For the purposes of this Regulation, the following definitions apply:

60. 'deep fake' means AI-generated or manipulated image, audio or video content that resembles existing persons, objects, places, entities or events and would falsely appear to a person to be authentic or truthful;

A very broad and vague definition:

- Is every AI-generated or manipulated media a deepfake?
- From whose perspective must the media appear authentic or truthful?



2. Transparency obligations under Art. 50 AI Act

(4) Deployers of an AI system that generates or manipulates image, audio or video content constituting a deepfake shall disclose that the content has been artificially generated or manipulated. [...]

⇒ **Adressing:** Deployer (D)

This image is generated by AI



In addition:

(2) Providers of AI systems, including general-purpose AI systems, generating synthetic audio, image, video or text content, shall ensure that the outputs of the AI system are marked in a machine-readable format and detectable as artificially generated or manipulated. [...]

⇒ **Adressing:** Provider (P)



3. Addition of a prohibition in Art. 5 AI Act

Amendment to the AI Act via the 'Omnibus on AI'

Debate over the addition of a prohibition on pornographic deepfakes

- The European Commission has not planned any further regulation of deepfakes
- Proposal by the European Council (13 March 2026):

(1) The following AI practices shall be prohibited:

ba) the placing on the market, the putting into service or the use of an AI system capable of generating, manipulating or reproducing realistic images, videos, audio or similar material of an identifiable natural person's intimate parts, or of an identifiable natural person engaged in sexually explicit activities, without that person's freely given, specific, informed, unambiguous and explicit consent for that generation, manipulation or reproduction;

⇒ **Addressing:** Provider (P), Deployer (D), User (A)

- Approval of the proposal by the European Parliament (26 March 2026)



III. Certification



III. Certification

- Are deepfake systems high-risk AI systems?
 - Decision against classifying them as high-risk AI systems during the legislative process
 - Consequently: **No mandatory certification** for deepfake AI systems
- **Voluntary certification** (Art. 95 AI Act)
 - with regard to the transparency obligations under Art. 50 (2) and (4) AI Act?
 - With regard to the planned prohibition under Article 5 (1) (ba) of the AI Act?
- Many unanswered questions regarding voluntary certification regarding transparency obligations and bans:
 - Who can be certified?
 - How can certification be obtained?
 - What are the benefits of such certification?





III. Certification

- One possible starting point for certifying the deepfake prohibition arises from Art. 5 (1a) AI Act, as set out in the European Council's proposal:

An AI system is capable of generating, manipulating or reproducing the content referred to where the system's design, training, architecture, capabilities or user-facing functionalities make that generation, manipulation or reproduction a reasonably foreseeable reproducible outcome, without requiring significant technical modification, and the system does not have effective technical safety measures and other safeguards to reliably prevent that generation, manipulation or reproduction and to reliably correct any observed or reported misuse.

- However, further potential starting points could also arise outside the AI Act, e.g. from media and press law
 - Duty of care under press law (section 19 MStV – State Media Treaty)
 - Example: AI generated Articles in the “Juristisches Internet Projekt Saarbrücken”



Thank you very much for your attention!



Prof. Dr. Georg Borges

georg.borges@uni-saarland.de | www.rechtsinformatik.saarland



Dipl.-Jur. Jonas Herrmann

jonas.herrmann@uni-saarland.de | www.rechtsinformatik.saarland

Further reading :

Borges, G.:
The European AI Act (AI Act) – Part 1: Overview, Scope and Initial Assessment, CR 2024, 497 ff.

Borges, G.:
The European AI Act (AI Act) – Part 2: Risk Management for High-Risk AI Systems, CR 2024, 565 ff.

Borges, G.:
The European AI Act (AI Act) Part 3 – Transparency Requirements, Enforcement, Overall Assessment, CR 2024, 633 ff.

