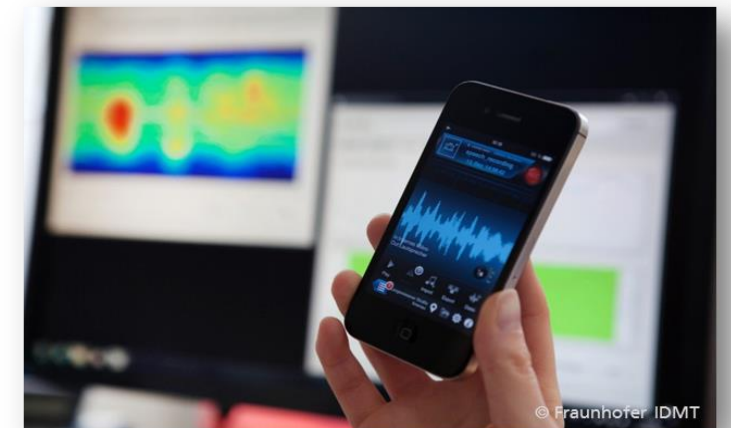


AUDIO-FORENSIK & CONTENT-VERIFIZIERUNG

Patrick Aichroth

Gruppenleiter Mediendistribution & Sicherheit, Fraunhofer IDMT

KAS Erfurt, 15.10.2020



Audioforensik für Content-Verifizierung

Überblick

■ Audioforensik

- Motivation und Themen Fraunhofer IDMT
- Ausgewählte Verfahren
- Zentrale Herausforderungen und das Projekt „DIGGER“
- Cheapfakes und Deepfakes

■ „Wahrheitssuche“

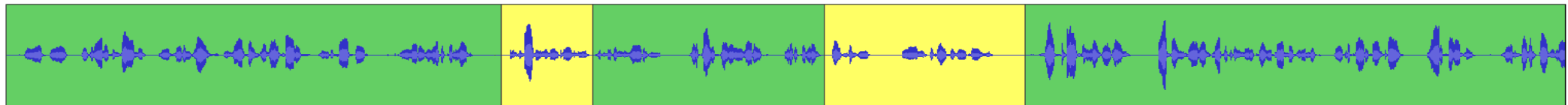
- Authentifizierung & Falsifizierung

■ Information und Desinformation

- Die Rolle von Technologie
- Probleme und Lösungsansätze

Audioforensik

Motivation

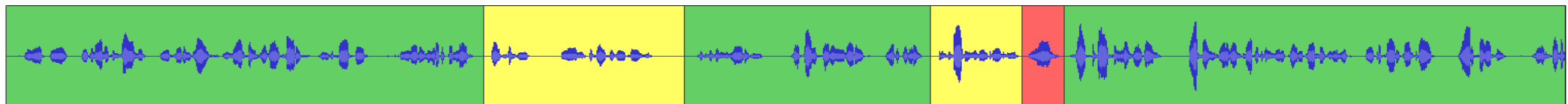


a economic environment that is growing, and one of the things we are very proud of is the fact that we had a 6%

contraction of the economy at the beginning of last year. This past quarter we had a 6%

increase in the growth of the economy.

That 12% swing offers greater opportunities for small businesses to prosper and thrive



a economic environment that is growing, and one of the things we are very proud of is the fact that we had a 6%

increase in the growth of the economy.

at the beginning of last year. This past quarter we had a 6%

contraction of the economy and

that 12% swing offers greater opportunities for small businesses to prosper and thrive

Audioforensik

Motivation

■ Treiber für „Shallow Fakes“:

- Manipulationen & Distribution: Einfach, schnell, günstig
- Immer mehr verfügbares Material
- Geringes Risiko der Entdeckung und “Bestrafung”
- Problemfälle nehmen zu:
Beweismittel, Verkaufsgespräche, Aufnahmen von Politikern / aus Krisengebieten, etc.
- Spoiler: auch künftig und für DeepFake Detection wichtig

■ Audio-Manipulationen (versus Bild/Video-Manipulationen)

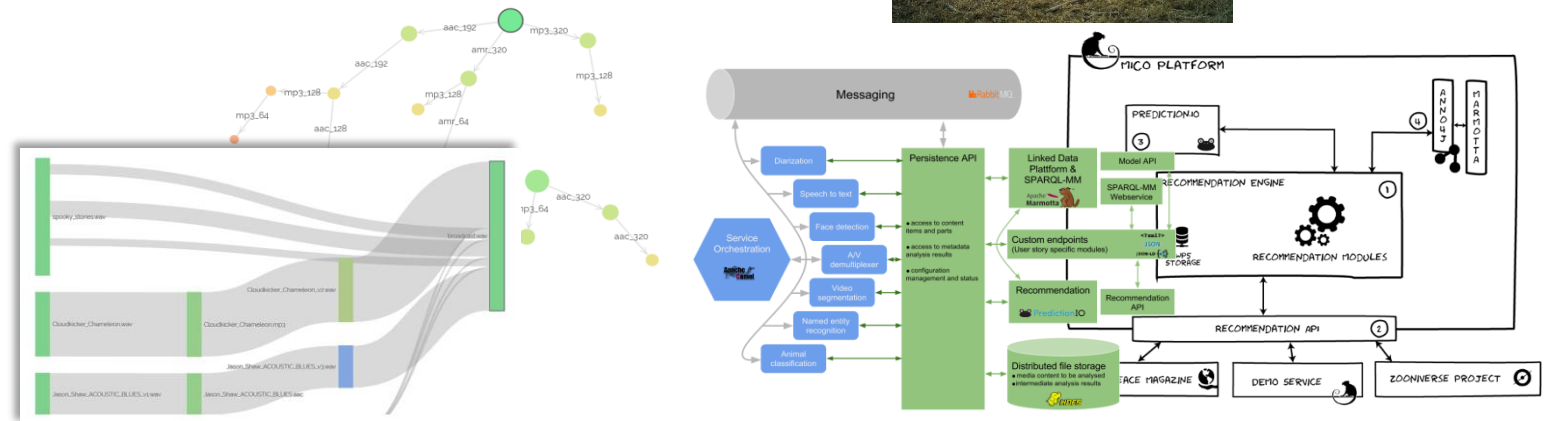
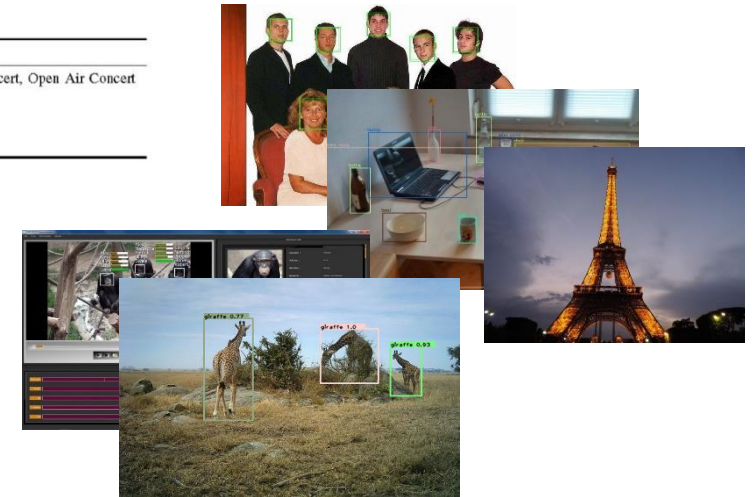
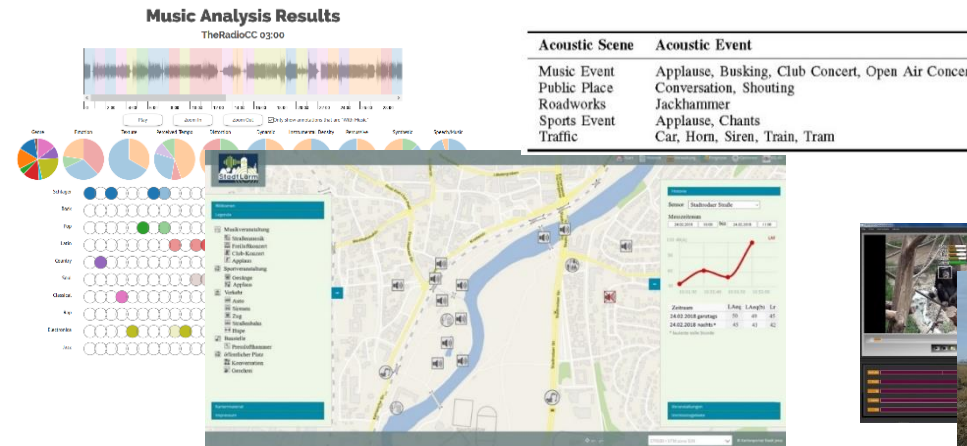
- wichtig, aber unterschätzt ;-)

Audioforensik (und andere Themen)

Fraunhofer IDMT: Abteilung Metadaten / M2D

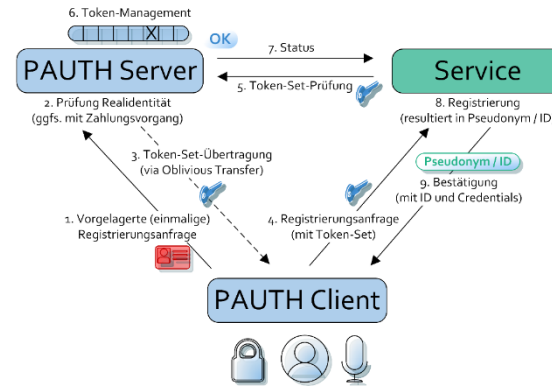
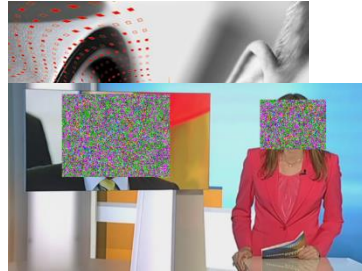
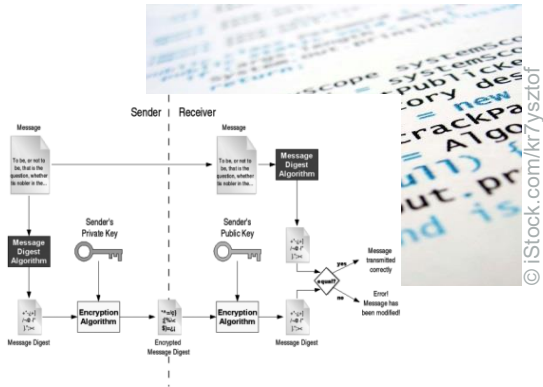
■ Content-Analyse und KI

- Audio- und Musikanalyse
- Visuelle Analyse
- A/V-Matching und Phylogenie-Analyse
- Multimodale Analyse
- Empfehlungssysteme

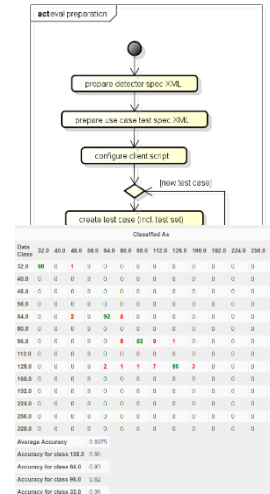
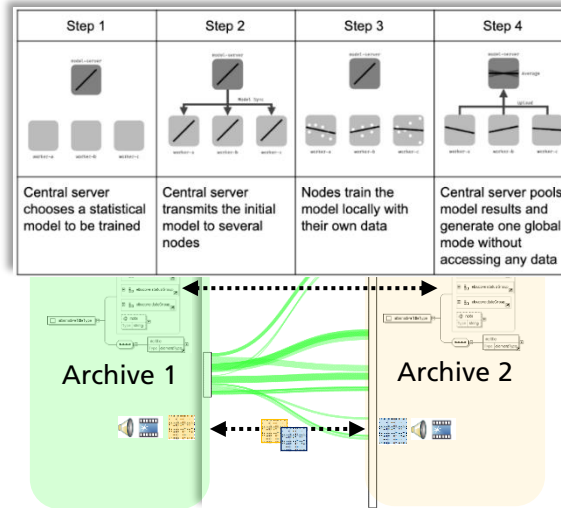


Audioforensik (und andere Themen)

Fraunhofer IDMT: Abteilung Metadaten / M2D



- Trust: Sicherheit, Datenschutz, Evaluation
- Security/Privacy-by-Design und PET
- Audioforensik und Medienauthentifizierung
- Dezentralisierung und On-Prem Analyse
- Systematische Evaluation

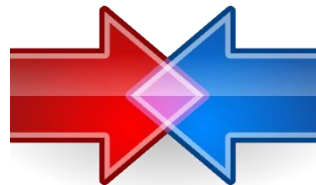


Audioforensik (und andere Themen)

Fraunhofer IDMT: Abteilung Metadaten / M2D

■ Content-Analyse und KI

- Audio- und Musikanalyse
- Visuelle Analyse
- A/V-Matching und Phylogenie-Analyse
- Multimodale Analyse
- Empfehlungssysteme

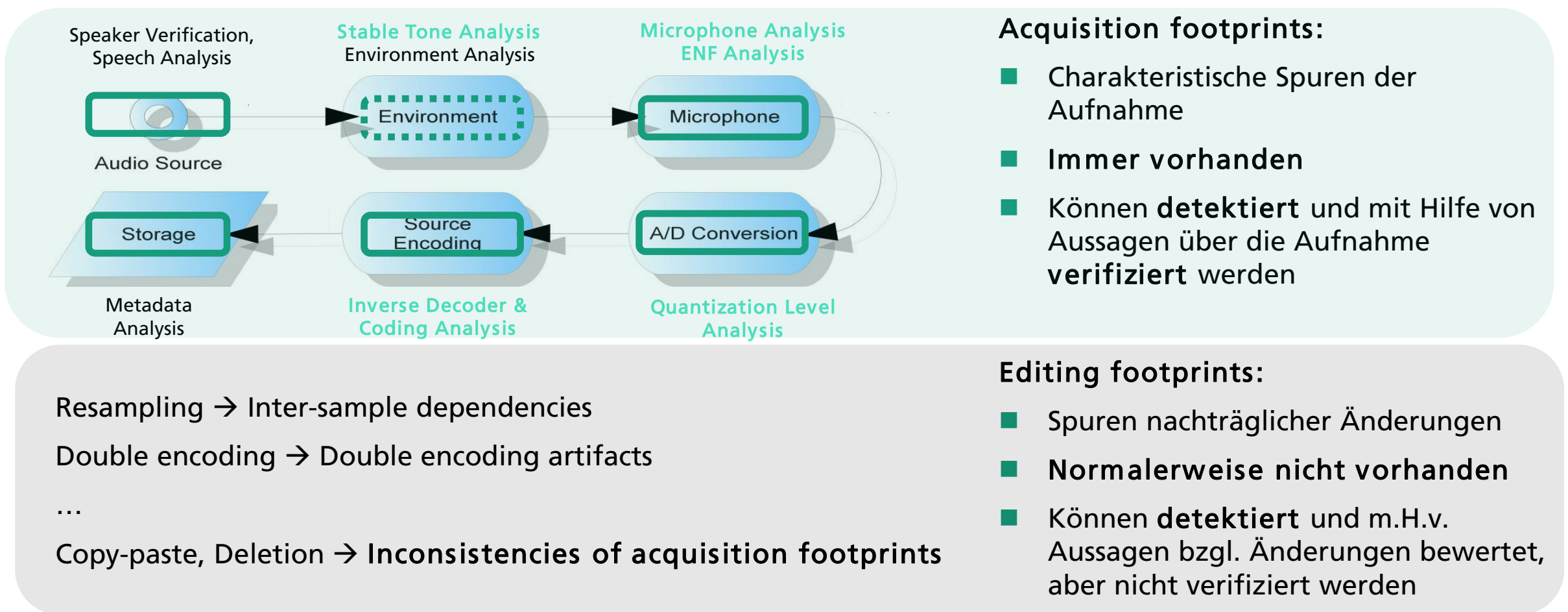


■ Trust: Sicherheit, Datenschutz, Evaluation

- Security/Privacy-by-Design und PET
- Audioforensik und Medienauthentifizierung
- Dezentralisierung und On-Prem Analyse
- Systematische Evaluation

Audioforensik

Verfahren: "Spurensuche" mit Detektoren



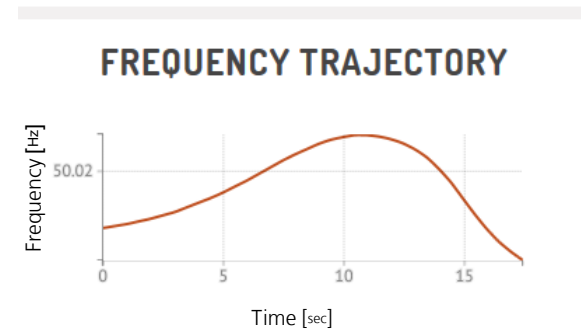
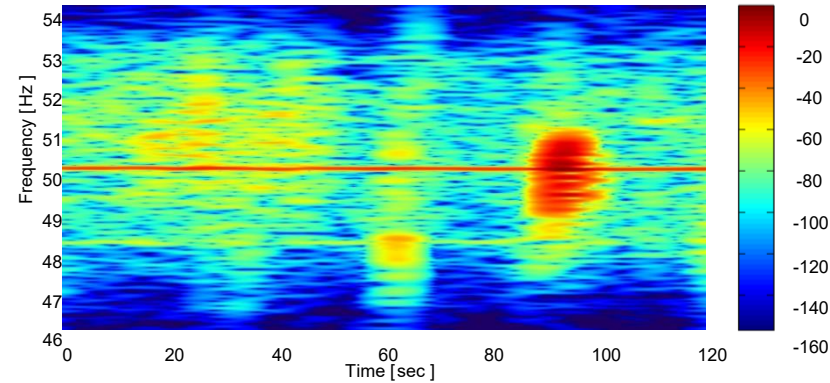
Audioforensik

ENF-Analyse

■ Analyse von "ENF Spuren"

- ENF: charakteristische Schwankungen um 50/60 Hz
- ENF Verlauf kann in Referenz-DB gespeichert werden

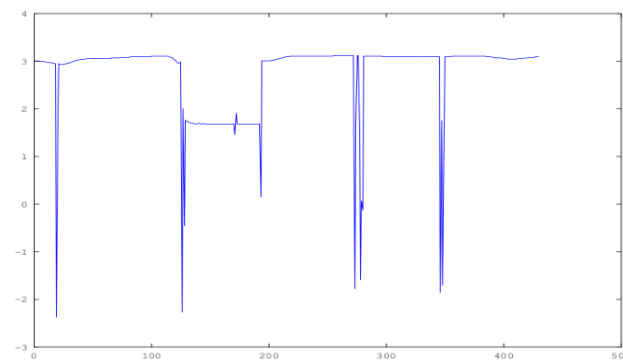
■ Einsatz:



Temporal Pattern Matching

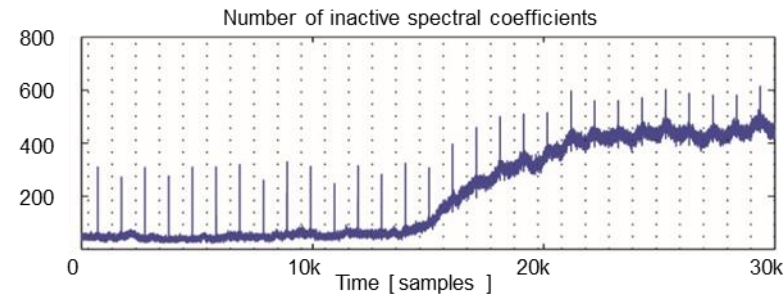
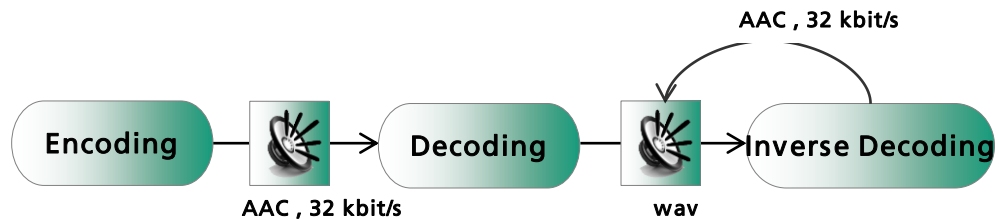


Detecting Discontinuities



Audioforensik

Inverse Decoder / Coding-Analyse



ENCODING INFORMATION	
Segment number:	3
Start:	00:00:05.400
Duration:	00:00:05.600
Codec:	MP3
Bitrate:	48

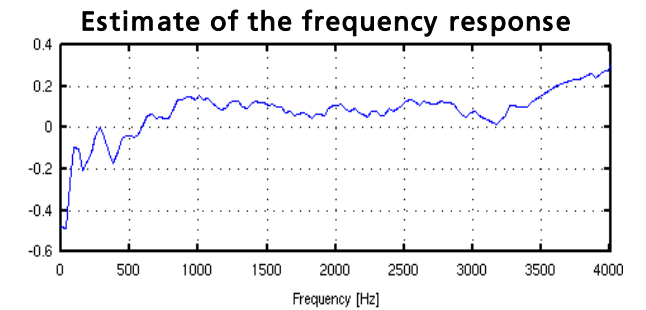
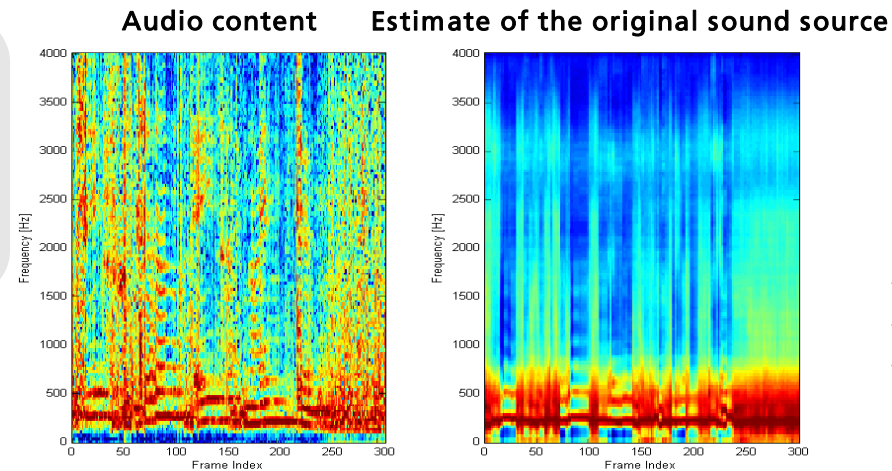
- Analyse von Codierungsspuren (e.g. MP3, AAC, GSM, ...)
 - "Reverse engineering" des Codierprozesses bzw. Multiple Coding Detection via ML
- Einsatz:
 - Erkennung von unterschiedlich encodiertem Ausgangsmaterial
 - Schnitt / Editieren in einer Aufnahme ← Inkonsistente "framing grid"
 - Erkennung von Mehrfachcodierung

Audioforensik

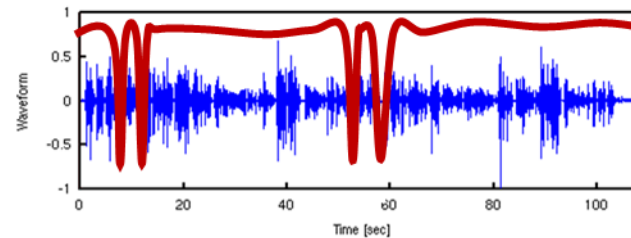
Mikrofon-Analyse

- Analyse des Aufnahmegeräts
 - Source/channel Estimation (ML)
 - → Frequency Response

■ Einsatz:



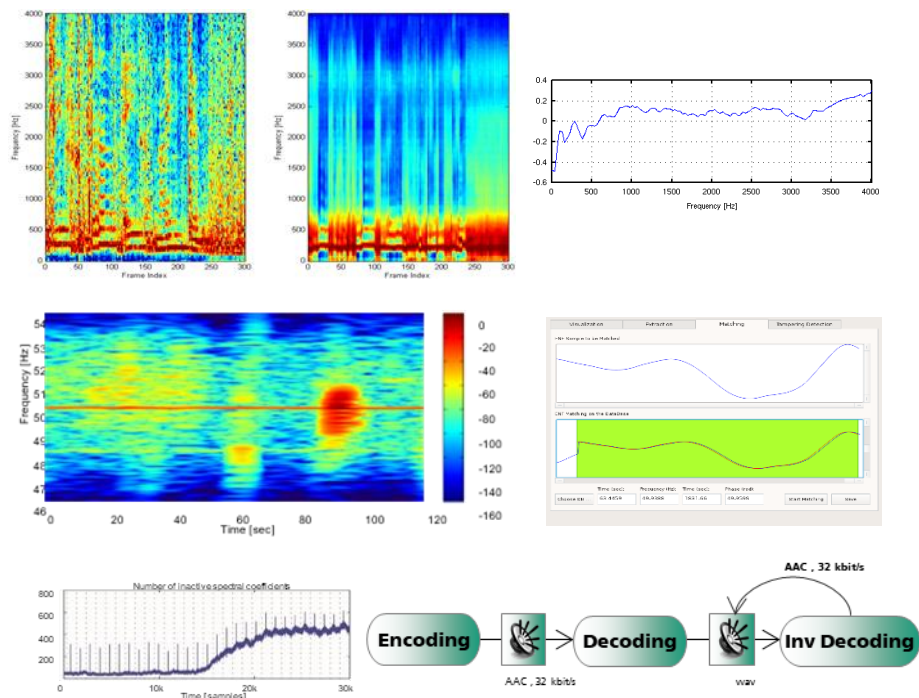
Microphone discrimination



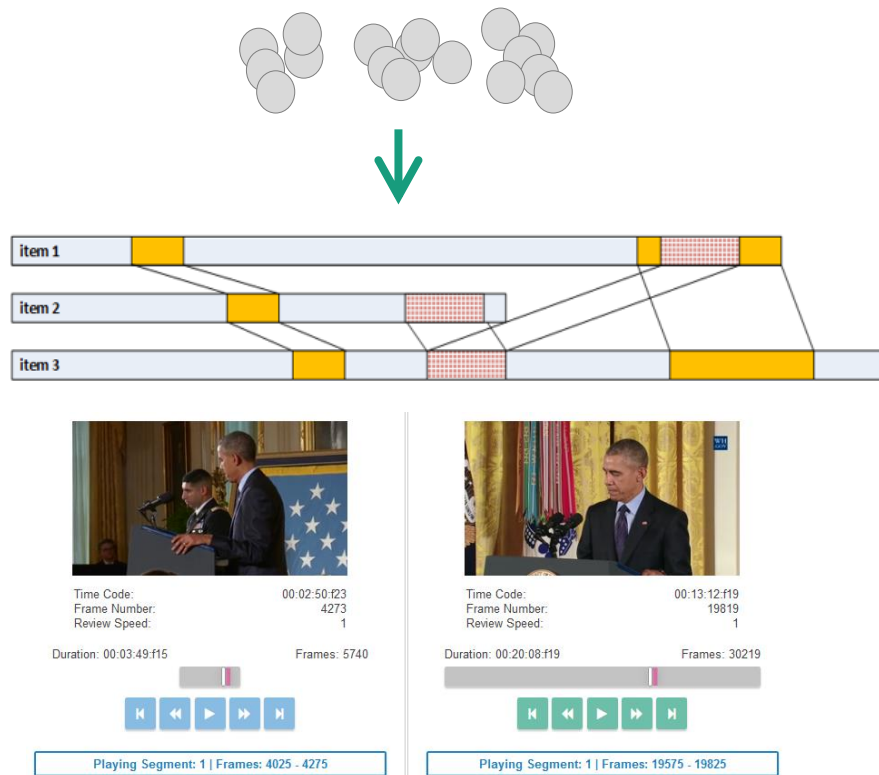
Audioforensik

Verfahren: zusätzliche Werkzeuge

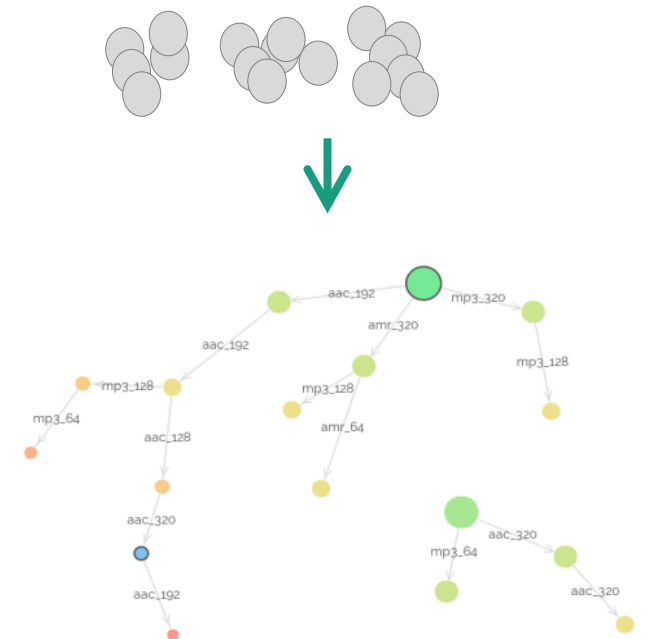
Audio Forensics Tools / Detectors



Audio Partial Matching



Audio Phylogeny Analysis

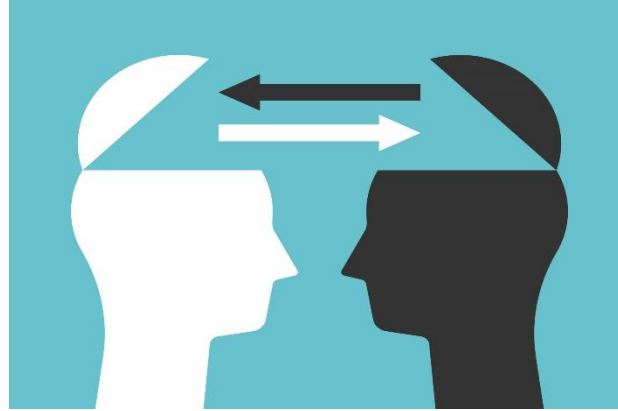


Audioforensik

Herausforderungen



© iStock.com/Serg_Velusceac



© iStock.com/Dmitrii_Guzhanin



“Können Sie mir sagen, ob dieses YouTube Video (URL) echt ist – ja / nein?
Ich habe leider **keine weiteren Infos** zur Herkunft.
Und ich bräuchte die Info **ASAP**, möglichst ohne **techn. Details** – Danke!”



Audioforensik für Journalismus

Truly Media → DIGGER

- <http://www.truly.media/>
- Plattform for Content-Verifizierung
- Fokus: Kollaborative, manuelle Analyse
- Enthält einige Video-Analyse-Werkzeuge

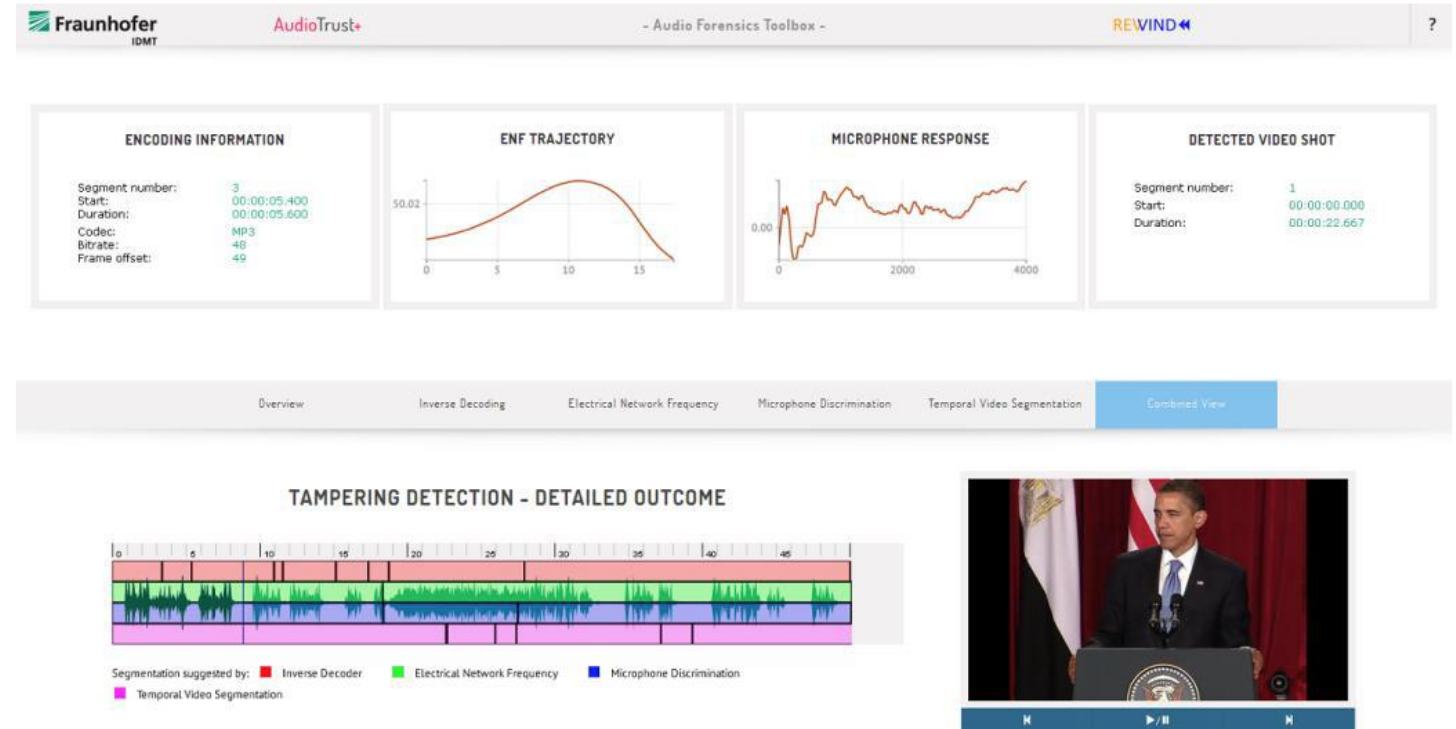


The screenshot displays the Truly Media DIGGER interface. On the left is a 'Checklist' sidebar with options like 'Add until date', 'Is geolocation possible?' (with 'Yes' and 'No' buttons), 'Location' (with a map and coordinates 38.717466, -90.199346), 'Landmarks/reference points', 'Latitude/Longitude', 'Preservation link', and 'Related Links'. Below this is a menu with 'Content', 'Source', 'Crediting & Permission', and 'Attachments'. The main area is titled 'Verification Tools' and shows a tweet from @realDonaldTrump. The tweet text is 'Sleepy Joe, in St. Louis, Missouri today: "We can only re-elect @realDonaldTrump"'. It includes a video thumbnail and engagement metrics (25610 retweets, 78787 likes). The interface also features a 'Multimedia' section with a video player showing a man speaking. At the top, there are tabs for 'Feeds' and 'Explore', and a 'Verification Status' bar with options: Raw, Pending, Unclear, Verified, Fake (selected), Graphic, License issue, and Profanity.

Audioforensik für Journalismus

Truly Media → DIGGER

- <https://digger-project.com/>
- Google DNI, 2019/07 - 2021/06
 - Deutsche Welle, ATC, IDMT
- Integration von Audioforensik-Werkzeugen
- Erkennung und Lokalisierung von A/V-Manipulationen
- Erkennung von synthetischem A/V-Material



Content-Verifizierung: Ausblick

Cheapfakes → deepfakes: Sprachsynthese

■ “DIGGER” Definitionen:

- **Cheapfakes:** **manipulated** A/V content created using ‘low tech’ like editing or speed adjustments which, changing context and/or meaning, can be very convincing
- **Deepfakes:** **synthetic** A/V content generated with technologies like **Deep** Learning which are **deeply** convincing / realistic

■ Quiz: Nur eine von vier Varianten A-D ist „natürlich“, drei synthetisch – welche?

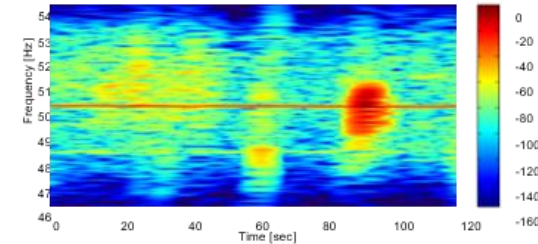
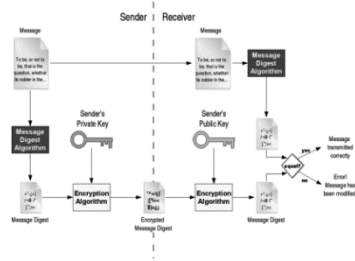
(vgl. Eric Battenberg et al.: Effective Use of Variational Embedding Capacity in Expressive End-to-End Speech Synthesis, <https://arxiv.org/abs/1906.03402>, vermutl. Wavenet-basiert)

	A	B	C	D
#1: <i>I've swallowed a pollywog.</i>				
#2: <i>Not in that way.</i>				
#3: <i>I charge you by all that is sacred, not to attempt concealment.</i>				

- Erkennung aktuell noch möglich, aber zunehmend schwierig (Architektur!) → F&E-Projekte!

Content-Verifizierung: Methoden / Grundlagen

Wahrheitssuche: Authentifizierung vs. Falsifizierung



Authentifizierung ("Beweis")

- Digitale Signaturen / Wasserzeichen
- Nachweis der "Echtheit"
- Jede Modifikation wird erkannt
- Generischer Ansatz
- Keine / kaum Details zu Veränderungen
- Erfordert Integration in / Modifikation der Aufnahme- und Verarbeitungsprozesse
- Eher die Ausnahme
- Gut für "proaktiven" Echtheitsnachweis

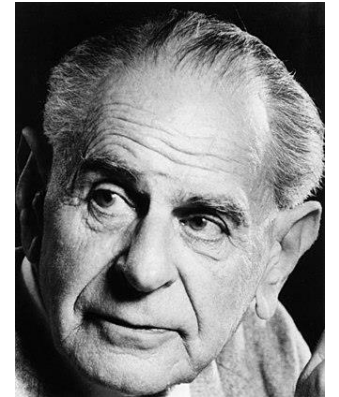
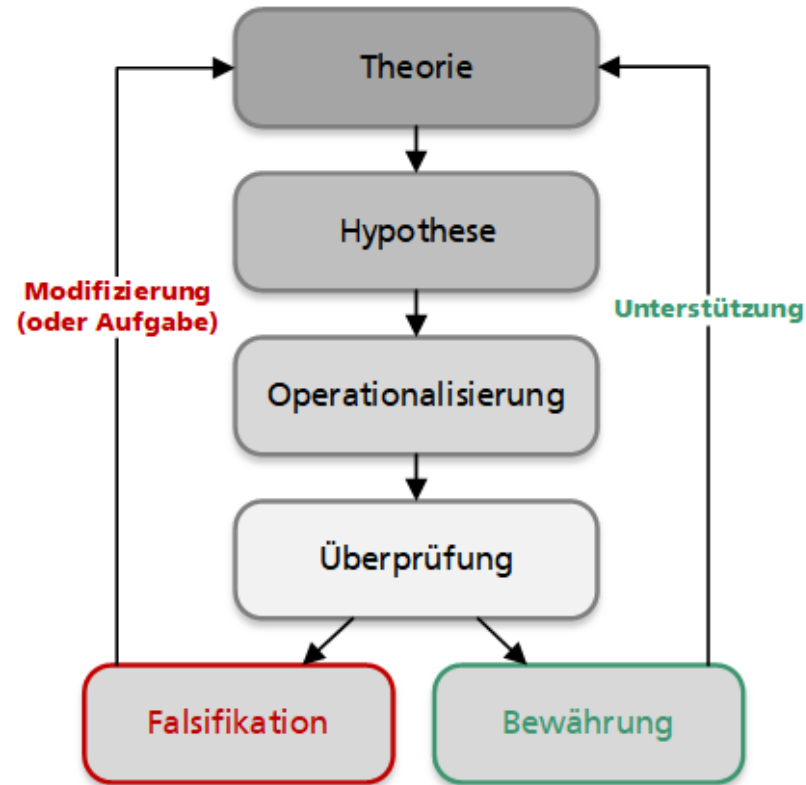
Tampering detection ("Falsifizierung")

- Signalanalyse und maschinelles Lernen
- Bestätigung / Falsifizierung von Aussagen
- Erkennung spezifischer Spuren
- Viele spezielle Verfahren
- Detailliertes Verständnis der Veränderungen
- Keine Veränderung der Aufnahme- und Verarbeitungsprozesse, aber Info erforderlich
- Eher die Regel
- Standard für die "passive" Prüfung

Content-Verifizierung: Methoden / Grundlagen

Wahrheitssuche: gemeinsame Basis Falsifizierung?

- Prüfung von Echtheit, aber auch **Richtigkeit**
- Berücksichtigung des **Kontextes**
- „maschinelle“ Prüfwerkzeuge können „menschliche“ Prüfung **unterstützen, aber nicht ersetzen**
- **Prüfprozess: starke Parallelen zu „kritischem Rationalismus“** (wiss. Erkenntnisprozess)
 - Gemeinsames „Wahrheitsverständnis“?
 - **Kritischer Rationalismus:** Wahrheit als regulatives Ziel, Falsifizierung als Weg zur Annäherung
 - **Postmoderne / Konstruktivismus:** es gibt nur konkurrierende subjektive Wahrheiten
 - Alles wird zur Machtfrage
 - Argumentationsbasis? Alternative Fakten? Personalisierung statt Argumente?



Information und Desinformation

Einige mögliche Lösungsansätze

■ Technologie:

- Authentifizierung, Manipulationserkennung
- „Tracking“ von Quellen und Verbreitungswegen
- „Anti-Echokammer“-Empfehlungsmaschinen

■ Normen:

- Falsifizierung als „gemeinsame Sprache“?
- Bewusstsein und Umgang mit „Human Biases“

■ Markt:

- Förderung neuer Geschäftsmodelle
- „Trustworthy AI“ (bei richtiger Umsetzung):
Chance für Europa und Deutschland?

■ Gesetze / Politik:

- Verbote von Technologien?? → vgl. OSS
- Dialog mit „Technologie“!
- Eher Tools als Regulation? Gefahren der „delegierten Kontrolle“ - wo sind die Grenzen?
- Förderung von Bildung und Training
- Forschung

AUDIO-FORENSIK & CONTENT-VERIZIFIZIERUNG

Vielen Dank!

Patrick Aichroth

Gruppenleiter Mediendistribution & Sicherheit, Fraunhofer IDMT

KAS Erfurt, 15.10.2020

