

03
Mai/Juni
2026

pvt

POLIZEI VERKEHR + TECHNIK

71. Jahrgang

€ 7,50

ISSN (Print) 0722-5962

www.pvtweb.de

Seit 1956 die Fachzeitschrift für Innere Sicherheit

Offizielle Publikation der GPEC[®] 2026



**Alles für die Innere Sicherheit,
Strafverfolgung und Heimatschutz**

20. – 22. Mai 2026, Leipzig

Vorbericht zur GPEC[®] 2026 Seite 19

Audioforensik für Sicherheitsbehörden

Analyse und Verifikation von Audioinhalten im Kontext von Shallowfakes und Deepfakes

Manipulierte Audioaufnahmen sind zunehmend Gegenstand realer Ermittlungs- und Betrugsfälle sowie journalistischer Recherchen. Audio stellt dabei einen zentralen Informationsträger dar: Insbesondere Sprache transportiert Aussagen, Intentionen und Identitätsmerkmale unmittelbar und macht Audioinhalte anfällig für Manipulationen. Bereits einfache Eingriffe, wie etwa das Entfernen, Umstellen oder gezielte Ersetzen einzelner Segmente, können die semantische Aussage einer Aufnahme erheblich verändern.

In der Ermittlungspraxis stellt sich daher häufig die Frage, ob eine Audioaufnahme als authentisch gelten kann, also ob sie mit dem ursprünglichen Aufnahme- und Entstehungskontext übereinstimmt oder nachträglich verändert wurde.

Angriffe auf Audiomaterial: Shallowfakes und Deepfakes

Grundsätzlich lassen sich zwei Angriffsklassen unterscheiden:

Shallowfakes entstehen durch Eingriffe in vorhandenes Audiomaterial, etwa durch Schneiden, Rekombination von Segmenten aus unterschiedlichen Aufnahmen oder gezielte Dekontextualisierung, bei der Aussagen bewusst in einen irreführenden Zusammenhang gestellt werden. Schon ein einzelnes kopiertes „Ja“ oder ein aus dem Kontext gerissener Satz können die Bedeutung einer Aufnahme völlig verändern. Solche Manipulationen sind vergleichsweise einfach realisierbar und dennoch für den Menschen kaum wahrnehmbar. In vielen praktischen Anwendungsszenarien stellen sie deshalb nach wie vor ein erhebliches Problem dar, und zu ihrer

Detektion werden insbesondere Verfahren zur Analyse von Konsistenz, Kontext und aufnahmespezifischen Charakteristika eingesetzt.

Deepfakes hingegen basieren auf KI-gestützten generativen Verfahren zur Synthese von Audioinhalten. Sie ermöglichen es, realistisch klingende künstliche Stimmen und Umgebungsg Geräusche zu erzeugen und gewinnen insbesondere im Kontext von Identitätsmissbrauch, Betrug und Desinformation zunehmend an Bedeutung. Die Entwicklung entsprechender Detektoren konzentriert sich hier vor allem auf die Erkennung von Synthesespuren sowie auf das Fehlen charakteristischer Merkmale natürlicher Aufnahmen.

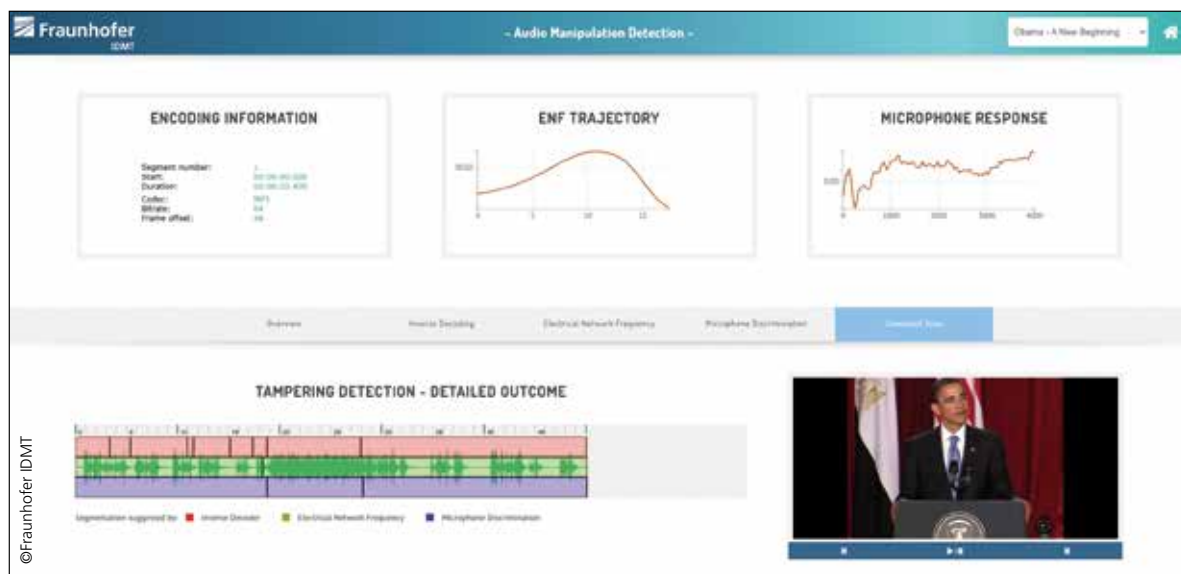
Methoden der Verifizierung und Authentifizierung

Zur Erkennung der oben genannten Angriffe kommen grundsätzlich verschiedene *aktive* und *passive* Verfahren zum Einsatz:

Aktive Verfahren nutzen Mechanismen wie digitale Signaturen oder Wasserzeichen, die bereits während der Aufnahme oder Verarbeitung eingebettet werden, um die Integrität von Audioinhalten sicherzustellen

und nachträgliche Veränderungen zuverlässig nachzuweisen. Ihr Einsatz ist jedoch auf kontrollierte Umgebungen beschränkt, in denen Entstehungs- und Verarbeitungsprozesse nachvollziehbar dokumentiert werden können. In allen anderen Fällen sind solche Ansätze nicht anwendbar. Außerdem erlauben aktive Verfahren keine detaillierten Aussagen über Art und Ausführung eines Angriffs.

Passive Verfahren setzen keine vorbereitenden Maßnahmen in kontrollierten Umgebungen voraus, müssen aber aufwändig für eine möglichst breite Palette von individuellen Angriffsfällen entwickelt werden. Sie basieren auf der Analyse charakteristischer Spuren, sog. „Footprints“, die bei Aufnahme, Kodierung und Weiterverarbeitung im Signal entstehen. Dazu zählen beispielsweise Verfahren zur Umgebungs- und Szenenklassifikation, zur Mikrofon- und Geräteklassifikation oder zur Analyse von Kodierungs- und Verarbeitungsspuren. Sie können beispielsweise zur Überprüfung von Hypothesen zum Aufnahmezeitpunkt, Aufnahmeort und zu Verarbeitungsprozessen sowie zur Detektion und zeitlichen Lokalisierung von Inkonsistenzen im Signal genutzt werden.



Die Analyse charakteristischer Signalspuren („Footprints“) wie zum Beispiel Kodierung, elektrische Netzfrequenz oder Mikrofonklassifizierung ermöglicht Rückschlüsse auf Herkunft, Bearbeitung und Konsistenz von Audioinhalten sowie die Detektion und zeitliche Einordnung möglicher Manipulationen.

Auf diese Weise können belastbare Hinweise auf Manipulationen oder inkonsistente Aussagen über den Aufnahmekontext ermittelt werden.

Während diese Aspekte weiterhin relevant bleiben, entstehen durch die mit erheblichem Ressourceneinsatz vorangetriebene Entwicklung generativer KI neue Herausforderungen: Verfahren müssen über verschiedene Stimmen und Syntheseverfahren hinweg robust generalisieren und zugleich nachvollziehbare, forensisch interpretierbare und im sicherheitskritischen Umfeld belastbare Ergebnisse liefern. Diese Anforderungen werden von bestehenden Detektionsverfahren bislang nur unzureichend erfüllt. Entsprechend kommt der Entwicklung verbesserter Ansätze in aktuellen Forschungsvorhaben besondere Bedeutung zu.

Neben Manipulationserkennung und Synthesedetektion gewinnt zudem die Analyse von Herkunft, Wiederverwendung und Verbreitung von Inhalten im Sinne einer Provenienz- bzw. Herkunftsanalyse an Bedeutung: Da Angriffe häufig auf bereits existierendem Material beruhen, ist die schnelle Erkennung der Wiederverwendung auch kurzer Segmente in Datensätzen entscheidend. Hierfür kommen sogenannte „Reuse-Detection-Verfahren“ zum Einsatz. Darauf aufbauend unterstützt die sogenannte „Phylogenie-Analyse“ die Rekonstruktion von Bearbeitungs- und Verbreitungsketten sowie die Einordnung von Inhalten in ihren zeitlichen und kontextuellen Zusammenhang. Damit lassen sich Entstehungsprozesse und Verbreitungsdynamiken nachvollziehen und für die Erkennung von Desinformationskampagnen nutzen.

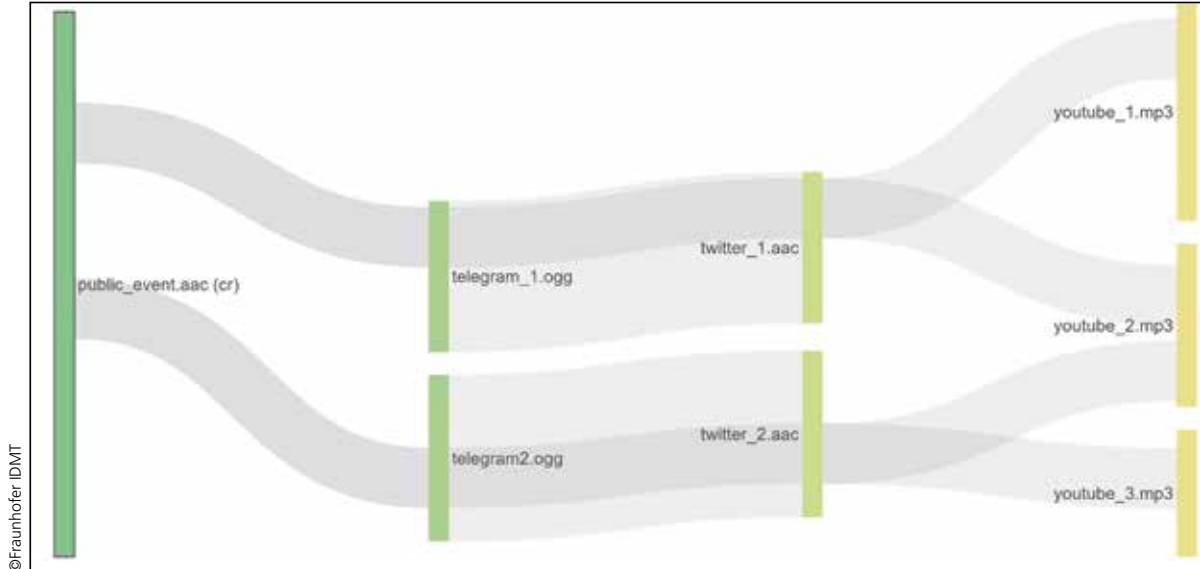
Der EU SOCTA 2025 Bericht von Europol unterstreicht die wachsende Bedeutung solcher Technologien, indem er die zunehmende Nutzung synthetischer Medien und Desinformation in hybriden Bedrohungsszenarien und Desinformationskampagnen beschreibt.

Anwendungsfelder und aktuelle Forschungsprojekte in der Audioforensik

Die beschriebenen Technologien wurden bereits in realen Analyseprozessen eingesetzt, etwa in der Medienverifikation sowie

im Kontext von Ermittlungen und Strafverfahren. Hierfür hat das Fraunhofer IDMT audiodenstische Werkzeuge entwickelt, die zur Analyse potenziell manipulierter Beweismittel, zur Verifikation von Inhalten, zur Untersuchung von Desinformationskampagnen sowie im Kontext sprachbasierter Betrugsformen eingesetzt werden.

zur Prüfung konkreter Hypothesen zum Aufnahme- und Entstehungskontext – von der Analyse aufnahmespezifischer Spuren über die Erkennung synthetischer Inhalte bis hin zur Provenienz- und Kontextanalyse. So entstehen belastbare, kontextualisierte Bewertungen, die über einzelne Indikatoren hinausgehen.



Herkunftsanalyse von Audiodateien: Das Fraunhofer IDMT arbeitet an Verfahren, die eine Rückverfolgung von Ursprüngen und Verbreitungswegen von Audiodateien oder Teilen von Audiodateien ermöglichen.

Die aktuellen Herausforderungen, insbesondere im Bereich der Synthesedetektion, werden in aktuellen Forschungsprojekten adressiert. Mit dem Projekt *SpeechTrust+* wurden am Fraunhofer IDMT Grundlagen zur Erkennung KI-basierter Sprachsynthese und Stimmverfremdung für die Beweismittelprüfung sowie zum Schutz vor Betrug und Desinformation geschaffen – in Zusammenarbeit mit den Landeskriminalämtern Baden-Württemberg und Bayern. Aufbauend darauf werden im Forschungsprojekt *PADSE* seit Kurzem neue Verfahren zur personenzentrierten Erkennung von Audiomanipulation und -synthese entwickelt, die insbesondere hinsichtlich Generalisierbarkeit und Erklärbarkeit deutliche Fortschritte gegenüber bestehenden Detektionsverfahren erwarten lassen. Darüber hinaus sind Projekte zur Weiterentwicklung der Herkunftsanalyse sowie weiterer forensischer Analyseverfahren geplant.

Fazit: Kombinierte Verfahren und enger Praxisbezug als Schlüssel für den erfolgreichen Einsatz

Aus den dargestellten Entwicklungen ergeben sich Implikationen für den praktischen Einsatz in Sicherheitsbehörden: Isolierte Detektionsansätze reichen nicht aus. Entscheidend ist die Kombination komplementärer Verfahren

Voraussetzung hierfür ist eine kontinuierliche Weiterentwicklung der Werkzeuge und eine systematische technische Evaluation unter realistischen Einsatzbedingungen. Dafür ist es wichtig, den Austausch zwischen Forschung und Anwendung weiter zu intensivieren, um den aktuellen Herausforderungen wirksam zu begegnen.

Zur Medienforensik-Webseite:



Fachlicher Ansprechpartner:

Patrick Aichroth

E-Mail: patrick.aichroth@idmt.fraunhofer.de

Telefon: +49 3677 467-121



Besuchen Sie uns zur GPEC® vom 20. bis 22. Mai 2026 in Leipzig auf dem Fraunhofer-Gemeinschaftsstand J32.