

## Pressemitteilung

Fraunhofer-Institut für Digitale Medientechnologie IDMT

Christian Colmer

05.09.2024

<http://idw-online.de/de/news839248>

Forschungs- / Wissenstransfer, Forschungsergebnisse  
Gesellschaft, Informationstechnik, Medien- und Kommunikationswissenschaften, Politik, Wirtschaft  
überregional



## So wird Medienpräsenz messbar: Software ermöglicht Medienanalysen durch kombinierte Gesichts- und Sprechererkennung

**IBC 2024: Softwarelösung »InsightPersona« des Fraunhofer IDMT ermöglicht aussagekräftige Medienanalysen durch eine kombinierte Gesichts- und Sprechererkennung** Auf der International Broadcasting Convention (IBC) in Amsterdam stellt das Fraunhofer-Institut für Digitale Medientechnologie IDMT eine Lösung vor, die KI-basiert Gesichter und Stimmen in großen Medienarchiven identifiziert. Die Software »InsightPersona« liefert in Sekundenschnelle Auswertungen zur Präsenz von Personen, Personengruppen und Gesprächsinhalten. In Verbindung mit Angaben zu Sprachverständlichkeit oder Aufgeregtheit in der Stimme bieten sich zudem neue Perspektiven für Suchen in Videoarchiven und Datenbanken.

05. September 2024. Wie oft ist der Bundeskanzler auf welchen Sendern sichtbar? Welche Worte fallen bei seinen Auftritten besonders häufig und wie aufgeregt klingt seine Stimme dabei? Wie häufig sind Frauen in einer bestimmten Fernsehsendung zu sehen – und wie groß ist ihr Sprechanteil? Diese und weitere Fragen zu Medieninhalten werden in Rundfunkanstalten, von Streaming-Plattformen, Redaktionen oder auch Forschungseinrichtungen gestellt. Sei es zur Optimierung der Content-Strategie, zur Recherche von O-Tönen und Videomaterial für den nächsten Nachrichtenbeitrag oder auch zur Analyse der Diversität im Fernsehprogramm. Mit InsightPersona hat das Fraunhofer IDMT dafür die passende Lösung entwickelt.

InsightPersona: Präzise Analyseergebnisse für Gesichter, Stimmen und wahrgenommene Geschlechter

Die KI-Algorithmen zur kombinierten Gesichts- und Sprechererkennung ermöglichen die schnelle Auswertung großer Archive mit vielen tausend Stunden Material hinsichtlich bestimmter Gesichter, Stimmen, wahrgenommener Geschlechter (männlich, weiblich) und der Unterscheidung Erwachsener oder Kind. Die schnelle Auswertung wiederum ermöglicht es, nahezu in Echtzeit zum Beispiel Nachrichtenbeiträge mit Audio- und Videoausschnitten passender Personen zusammenzustellen.

Die Ergebnisse der Suchaufträge werden kundenindividuell visualisiert, wie zum Beispiel durch Heatmaps oder Zeitleisten. »Diese Kombination aus audiobasierten und visuellen Erkennungstechnologien ermöglicht uns eine hohe Aussagekraft und Qualität der Suchergebnisse. Das ist besonders dann hilfreich, wenn sich die gewonnen Informationen gegenseitig ergänzen, beispielsweise wenn Personen in einem Medienbeitrag sprechen, aber gerade nicht im Bild gezeigt werden«, erläutert Uwe Kühhirt, Experte für Videoanalyse am Fraunhofer IDMT.

Einen Beitrag zur Barrierefreiheit von Medieninhalten kann die Anzeige von Sprachverständlichkeit liefern. Mediatheken, Streaming-Dienste oder Kommunikationsdienstleister könnten ihren Kunden damit Mehrwerte bieten, wie zum Beispiel durch die Auswahl besonders gut verständlicher Beiträge oder das Angebot einer alternativen Tonspur mit optimierter Verständlichkeit.

Weitere Features: Visualisierung von Gesprächsinhalten, Aufgeregtheit und schnelle Gesichtserkennung in großen Archiven

Zur diesjährigen IBC in Amsterdam präsentieren die Expertinnen und Experten des Fraunhofer IDMT ihre Softwarelösung live am Messestand des Fraunhofer-Geschäftsbereichs Digital Media. Besucherinnen und Besucher können dabei auch folgende Features kennenlernen:

- Für einzelne Beiträge aus Medienarchiven können Wortwolken angezeigt werden. Damit werden wichtige Schlüsselbegriffe der Gesprächsinhalte visuell dargestellt. So kann beispielsweise der inhaltliche Tonus einer politischen Ansprache auf einen Blick erfasst werden.
- Mit Hilfe einer farbigen Markierung von Tonspuren wird das Level an Aufgeregtheit in der Stimme, das sogenannte Excitement, visualisiert. »Spricht eine Person sehr ruhig und gefasst, ist die Tonspur grün gefärbt. Ist die Stimmfarbe sehr aufgeregt, wird sie gelb oder rot«, erklärt Christian Rollwage, Leiter der Gruppe Audiosignalverbesserung am Fraunhofer IDMT.
- Mit der InsightPersona Instant Face Search können außerdem große Videosammlungen sehr schnell nach beliebigen Gesichtern durchsucht werden. Alles, was dafür notwendig ist, ist ein Referenzbild mit dem Gesicht der gesuchten Person. Das ermöglicht eine schnelle Suche nach bestimmten Personen in einem bestimmten Datenset.

Eine passgenaue Lösung für jeden Kunden

Die Softwarelösung InsightPersona kann lokal beim Kunden (On-Premises) oder bei einem Dienstleister (Off-Premises) ausgeführt werden. Auch das Fraunhofer IDMT bietet seinen Kunden die Auswertung von Medienarchiven mit InsightPersona an. Das Dashboard mit seinen Analysemodulen und Statistiken kann durch das Fraunhofer IDMT auf die Auswertungsbedarfe des Kunden angepasst werden. Es stehen verschiedene Lizenzierungsmodelle für InsightPersona zur Verfügung.

Lernen Sie »InsightPersona« kennen und besuchen Sie uns vom 13. bis 16. September 2024 auf der IBC in Halle 8 am Stand B.80 des Fraunhofer Geschäftsbereichs Digital Media. Unsere Expertinnen und Experten besprechen gerne mit Ihnen die Vorteile dieses crossmodalen Analysetools für Ihre individuellen Einsatzbereiche.

Über das Fraunhofer IDMT

Das Fraunhofer-Institut für Digitale Medientechnologie IDMT ist eines von 76 Instituten und Forschungseinrichtungen der Fraunhofer-Gesellschaft, eine der führenden Organisation für anwendungsorientierte Forschung.

Am Hauptsitz des Fraunhofer IDMT in Ilmenau (Thüringen) arbeiten die Expertinnen und Experten an der sicheren und effizienten KI-basierten Erkennung und Klassifizierung von Audio- und Videodaten. Anwendungsbereiche sind Medien, industrielle Fertigung, Verkehr und Logistik sowie Umwelt und Agrarwirtschaft. Ein weiterer Schwerpunkt ist die Entwicklung von maßgeschneiderten Lösungen zur Produktion und Wiedergabe authentischer und räumlicher Klangerlebnisse für die Bereiche Professional Audio, Entertainment und Automotive.

Der Oldenburger Institutsteil Hör-, Sprach- und Audiotechnologie HSA steht für marktnahe Forschung und Entwicklung mit Schwerpunkten auf Sprach- und Ereigniserkennung, Klangqualität und Sprachverständlichkeit sowie Mobile Neurotechnologie und Systeme für eine vernetzte Gesundheitsversorgung. Mit eigener Kompetenz in der Entwicklung von Hard- und Softwaresystemen für Audiosystemtechnologie und Signalverbesserung setzen die Mitarbeitenden am Standort Oldenburg wissenschaftliche Erkenntnisse in kundengerechte, praxisnahe Lösungen um.

Weitere Informationen auf

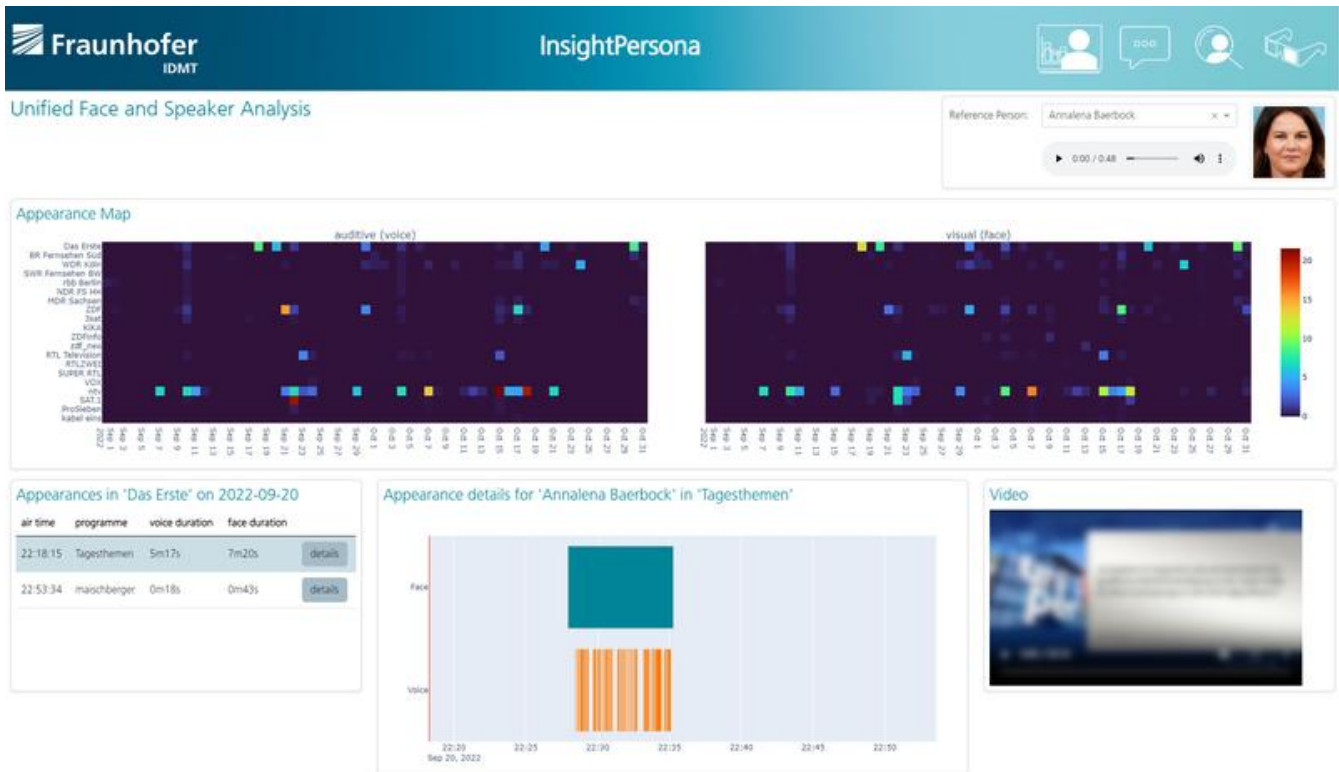
[www.idmt.fraunhofer.de](http://www.idmt.fraunhofer.de) und [www.idmt.fraunhofer.de/hsa](http://www.idmt.fraunhofer.de/hsa)

Kontakt für die Medien:  
Christian Colmer  
Leiter Marketing und Kommunikation  
Fraunhofer-Institut für Digitale Medientechnologie IDMT  
Institutsteil Hör-, Sprach- und Audiotechnologie HSA  
Marie-Curie-Str. 2  
26129 Oldenburg  
Telefon +49 441 2172-436  
[christian.colmer@idmt.fraunhofer.de](mailto:christian.colmer@idmt.fraunhofer.de)  
<http://www.idmt.fraunhofer.de/hsa>

URL zur Pressemitteilung: <https://www.idmt.fraunhofer.de/de/institute/projects-products/insightpersona.html>



Die KI-Software »InsightPersona« des Fraunhofer IDMT identifiziert in Sekundenschnelle Gesichter und Stimmen in großen Medienarchiven.  
[Fraunhofer IDMT/istock.com/vm](https://www.istock.com/vm)  
[Fraunhofer IDMT/istock.com/vm](https://www.istock.com/vm)



Die Ergebnisse der audiovisuellen Erkennung bestimmter Personen werden in einem leicht verständlichen und intuitiv bedienbaren Dashboard dargestellt und können für Trendanalysen und Statistiken verwendet werden.

Fraunhofer IDMT  
Fraunhofer IDMT