

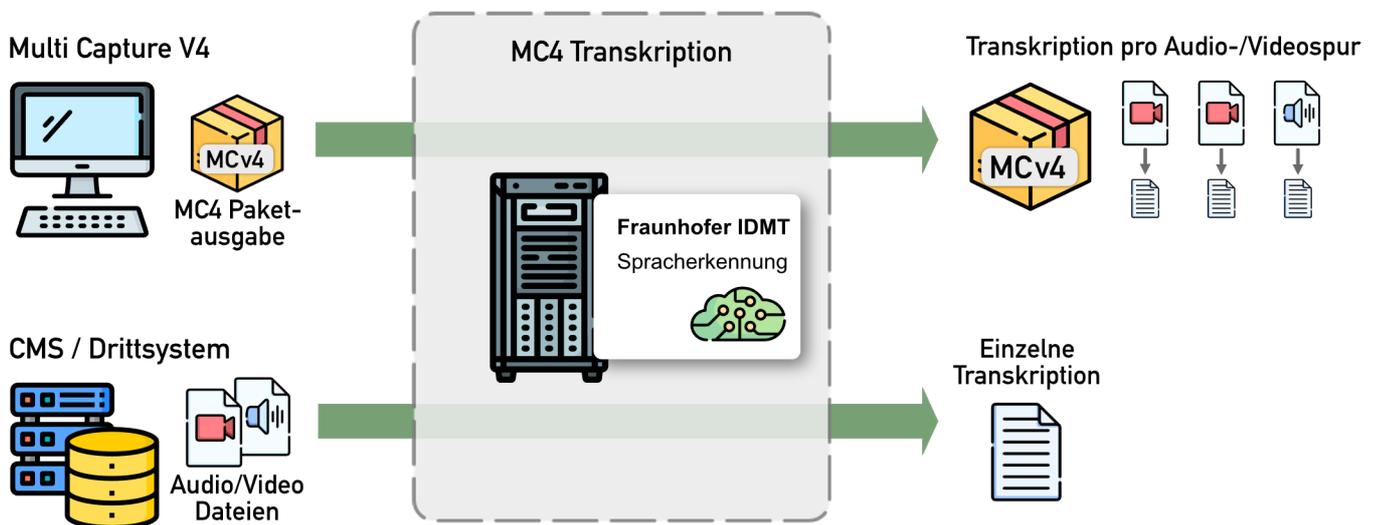
# Multi Capture Version

# 4

## Multi Capture V4 – Transkription

Die Multi Capture V4-Transkription ist eine serverbasierte Standalone-Lösung zur automatischen Verschriftung von Audio/Video-Daten. Die gesamte Verarbeitung findet ohne die Verbindung zu Drittsystemen statt, so dass die Daten niemals die eigene IT-Infrastruktur verlassen müssen. Der Datenaustausch und das Zwischenspeichern zu Verarbeitungszwecken erfolgt ausschließlich verschlüsselt (SSL / AES 256). Nach einer Transkription werden alle Nutzdaten auf dem MCv4 Transkriptionsserver rückstandslos gelöscht. Abhängig von der Qualität der Audio-/Videodateien und der verwendeten Hardware dauert die Verschriftung einer 30-minütigen Aufnahme ca. 3 Minuten und 20 Sekunden. Um eine zeitgerechte Verarbeitung zu gewährleisten, kann die Anzahl der Server gemäß der Menge und Länge des ankommenden Materials gewählt werden.

Die MCv4-Transkription arbeitet zusammen mit dem bekannten Multi Capture V4 Aufzeichnungssystem. Durch den Aufruf per Kommandozeile und entsprechenden Parametern können MCv4-Ausgabepakete um eine automatische Verschriftung der enthaltenen Audio- und Videospuren ergänzt werden. Für jede dieser Audio-/Videospuren wird dabei eine separate Text-Datei im Ausgabepaket hinzugefügt. Zudem bieten wir die Entwicklung von kundenspezifischen Anforderungen und Schnittstellen zur direkten Verschriftung von Audio/Video-Daten z. B. für Drittsysteme als maßgeschneiderte Lösung an



## Fraunhofer IDMT - Spracherkennung

Das Fraunhofer IDMT (Fraunhofer-Institut für Digitale Medientechnologie) liefert als Kooperationspartner die Komponenten für die interne Spracherkennung von MCv4-Transkription. Diese sind Ergebnisse langjähriger Forschung und Entwicklung im Bereich der computergestützten Spracherkennung und vereinen die neuesten Erkenntnisse mit aktuellen Technologien.

## Informationen zur Spracherkennung

Bei der Analyse mit anonymisierten Aufzeichnungen aus Polizeivernehmungen und Gerichtsverläufen erreichte die Spracherkennung bei Erwachsenen nach dem Stimbruch in deutscher Sprache eine Erkennungsrate bis zu 80%. Zum Vergleich wurden die zugehörigen manuellen Transkriptionen herangezogen. Derzeit können Sprachmodelle zur Erkennung von deutscher und englischer Sprache angeboten werden. Ein gesondertes Sprachmodell für Kinder befindet sich im Aufbau. Weitere spezialisierte Sprachmodelle für gesonderte Personengruppen oder andere Sprachen können auf Anfrage produziert und nachinstalliert werden.

Die Spracherkennung liefert gesprochenen Text mit zeitlichem Bezug zum verarbeiteten Audio-/Videomaterial. Durch geplante Weiterentwicklungen können zukünftig weitere Informationen automatisch erkannt werden (z.B. Sprechererkennung bei Aufnahmen mit mehreren Personen, automatische Markierung von Schlüsselwörtern oder Formulierungen). Bei Bedarf ist auch die Entwicklung einer Konvertierung in andere Formate möglich.

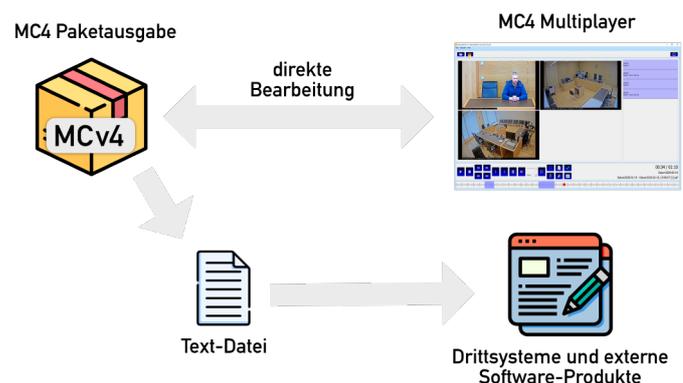
Allgemein stellt Spracherkennung besondere Anforderungen an die Qualität des Audio-Materials. Audio-Aufzeichnungen sollten optimaler Weise in CD-Qualität (44.1kHz - 16bit) erstellt werden, um die volle Bandbreite gesprochener Sprache wiedergeben zu können. Die MC4-Ausstattungen für mobilen oder stationären Einsatz erfüllen diese Anforderung.



## Unterstützung bei der anschließenden Kontrolle

Trotz fortgeschrittener Technik bei der Erkennung von gesprochener Sprache muss bei wortgenauen Texten (z.B. Vernehmungs- oder Verhandlungsverschriftungen) eine manuelle Kontrolle durch Menschen erfolgen. Eine automatische Verschriftung kann hier als Vorlage dienen und im Nachgang kontrolliert und korrigiert werden. Im Vergleich zur manuellen Vollverschriftung mit einem typischen Aufwand von 6/1 (6 Stunden Verschriftungsaufwand pro 1 Stunde Audio-Material) kann eine deutlich zeitgerechtere Bearbeitung erfolgen. Der MCv4-Multiplayer bietet hierzu weitere Unterstützung. Verschriftete MCv4-Ausgabepakete können direkt geöffnet und per Fußschalter (z.B. Fußschalter 540 von Grundig) gesteuert abgespielt werden. Der MC4-Multiplayer stellt den Text zeitsynchron zum Audio/Video dar und kann ihn im Editorbereich bearbeiten. Zur Integration in einen Arbeitsablauf kann der Multi-Player per Kommandozeile mit Parametern gestartet werden. Dadurch kann beispielsweise ein bestimmtes MCv4-Ausgabepaket per Mausklick aus einer Verwaltungsmaske heraus geöffnet werden.

Für andere Programme kann die Text-Datei herangezogen werden. Dort liegt der erkannte Text in reiner Form vor. Schreibkräfte können den erkannten Text einfach öffnen und in ihre gewohnte Arbeitsumgebung als Vorlage übernehmen.



## Betrieb

MCv4 Transkription ist für den Betrieb im eigenen Netzwerk konzipiert. Es werden keine Verbindungen zu externen Systemen oder Internet verwendet. MCv4-Transkription wird vorinstalliert auf einer Standard-Server-Hardware oder einem PC-System geliefert und stellt keine besonderen Anforderungen an die vorhandene IT-Struktur.

Update Juni 2021

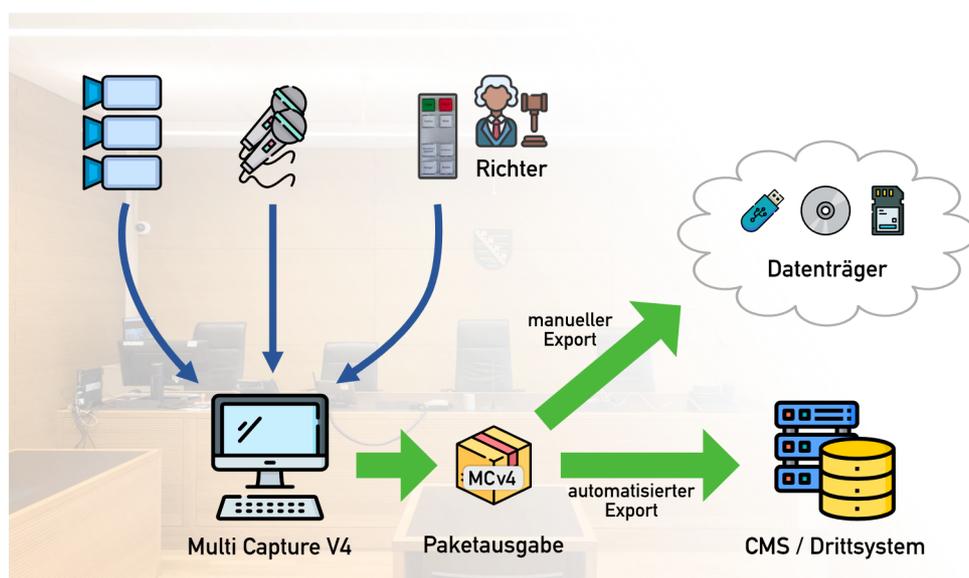
## Betrieb auf stationären oder mobilen Aufzeichnungssystem

Zum vereinfachten Betrieb und zur Verbesserung der Anwendungsfreundlichkeit für die Gerichte/Polizei wird ab Juni 2021 zusätzlich ein Lokal-Lizenz-Modell angeboten. Dadurch kann die Transkriptionsfunktion mit Einzelplatz-Lizenz direkt auf stationäre oder mobile Aufzeichnungssysteme (nach-)installiert werden. Die Transkription kann dadurch ohne Verbindung zu externen Systemen komplett offline verwendet werden. So können die Fließtexte zu Audio-/Videospuren schon während oder direkt nach der Aufzeichnung erstellt werden. Die Spracherkennung verwendet ebenfalls einen Fraunhofer-Sprachkern und liefert die gleiche oben beschriebene Qualität.

## Praxisbeispiel Gerichtsvernehmung und Verschriftung 1/2

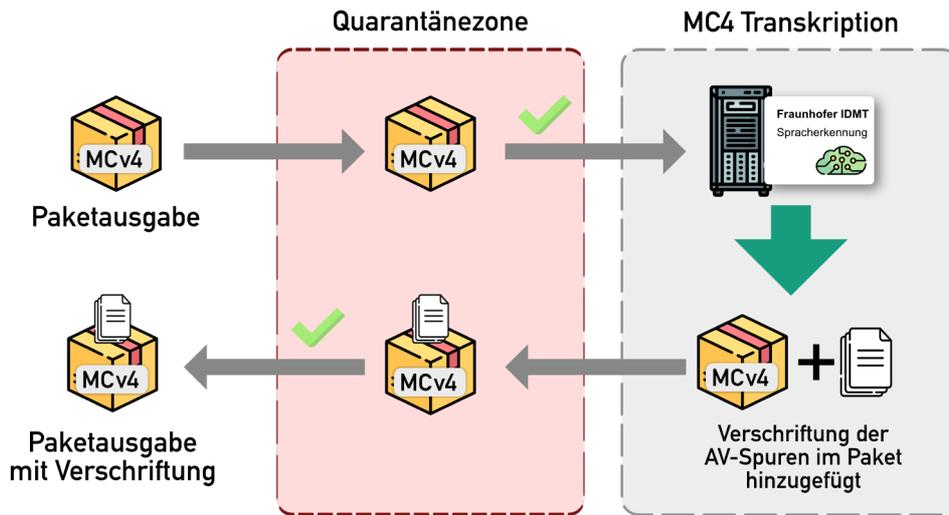
Im Folgenden wird exemplarisch der Einsatz von MC4 Transkription anhand einer vorhandenen stationären Multi Capture Anlage im Gerichtsraum dargestellt. Die Anlage besteht aus insgesamt 3 Kameras und 6 Raummikrofonen, die an ein PC-System mit Multi Capture V4 Software angeschlossen sind. Über eine Fernbedienung am Platz kann der Richter die Aufnahme für die Verhandlung per Tastendruck starten und stoppen. Zusätzlich bietet die Fernbedienung die Möglichkeit, weitere Programmfunktionen zu steuern (z.B. Konferenz- und Telefonzuschaltung, Kamerabild deaktivieren, Stummschaltung).

Nach Stoppen der Aufzeichnung kann die gesamte Aufzeichnung zur Archivierung als MC4-Ausgabepaket auf einen externen Datenträger DVD/BD, SD-Karte, USB-Stick oder Netzlaufwerk exportiert werden. Das MC4-Ausgabepaket ist eine definierte Dateistruktur als Zip-Archive, in der die Video- und Audiospuren zusammen mit allen Informationen der Aufzeichnungen organisiert sind.



## Praxisbeispiel Gerichtsvernehmung und Verschriftung 2/2

Ausgehend von einer Speicherung auf einem Netzlaufwerk kann eine automatische Transkription als nachgelagerter Prozess umgesetzt werden. Dazu wird das MC4-Paket in eine Quarantänezone exportiert, wo die Dateien vor der Weiterverarbeitung einer Sicherheitsüberprüfung unterzogen werden. Im Anschluss werden die MC4-Pakete durch MC4-Transkription verarbeitet. Die enthaltenen Video- und Audiospuren werden jeweils automatisch verschriftet und strukturiert dem MC4-Paket hinzugefügt. Nach erneuter Überprüfung in der Quarantänezone steht das verschriftete MC4-Paket für die manuelle Kontrolle bereit.



Mitarbeiter können gezielt einzelne Audiospuren zusammen mit der zugehörigen Verschriftung als Text-Datei in ihren gewohnten Arbeitsprozess übernehmen oder eine Kontrolle mittels MC4-Multiplayer durchführen.

