

# Thorsten Butz With the Beatles

#### **Community in motion 2024**

KI im Geschäftsumfeld – Möglichkeiten, Einsatz, Nutzen

https://cim-lingen.de

19. und 20. September 2024

IT-Zentrum Lingen

Vortrag: 19. September 2024, 15:00 Uhr

Seminarraum 2 - Use Cases

#### Links zu allen Videos, Demos, Folien und Codebeispielen:

```
https://www.thorsten-butz.de/cimlingen2024/
```



#### With the Beatles

Peter Jacksons Dokumentation "Get Back" nahm vorweg, was mit der Veröffentlichung der letzten Beatles-Single "Now and then" für jeden eindrucksvoll hörbar wurde: mit Hilfe künstlicher Intelligenz lassen sich heute Quantensprünge in der **Verarbeitung von Ton und Sprache** erzielen.

Besonders beeindruckend ist, dass zahlreiche Werkzeuge, die auf diesen neuen Technologien basieren, jedem zugänglich und ganz praktisch anwendbar sind - in vielen Fällen sogar kostenfrei. Mit Diensten wie Whisper von OpenAI oder der deutschen Plattform DeepL, lassen sich Transkripte und Übersetzungen erstellen, die Textqualität verbessern, Untertitel und Zusammenfassungen von längeren Texten erzeugen. Mit dem Anbieter Auphonic aus Österreich lassen sich mangelhafte Audiospuren aufwerten und vormals komplexe Prozesse mit Hilfe weniger Zeilen Scriptcode automatisieren.

Ich zeige am praktischen Beispiel, wie man mit Hilfe künstlicher Intelligenz und wenigen Zeilen PowerShell-Code vorhandenes Videomaterial

- klanglich optimiert: Lautstärke anpassen, Rauschen entfernen, Klang der Stimme aufwerten (Levelling, Normalization, Voice-EQ, Denoising)
- eine Mitschrift des gesprochenen Wortes erstellt
- eine Zusammenfassung der Inhalte generiert



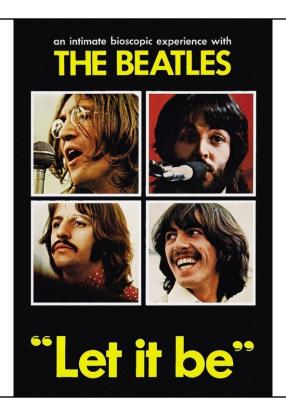




#### Im Fokus:

- https://auphonic.com
  - Auphonic ist ein 2003 von Georg Holzmann gegründeter Webdienst aus Graz, Österreich.
- <a href="https://openai.com/index/whisper/">https://openai.com/index/whisper/</a>
  - Whisper wurde 2022 quelloffen unter der MIT Lizenz veröffentlicht.
  - Der Dienst wurde unter dem Dach von openAl entwickelt und lässt sich in eigene Lösungen (lokal) integrieren. Whisper ist (als eigene Instanz) in Auphonic integriert.
- https://www.deepl.com
  - DeepL ist ein fortgeschrittener Übersetzungsdienst, der 2017 von der gleichnamigen Kölner Firma (DeepL SE) veröffentlicht wurde.

1969/70



**56 Stunden Film- und 140 Stunden Audiomaterial** entstanden im Januar 1969 unter dem Regisseur Michael Lindsay-Hogg. Aufgrund zahlreicher Unstimmigkeiten zwischen den Musikern, fand der Film keine große Verbreitung obgleich er 1970 im Kino gezeigt wurde.

(Die Länge des Spielfilms "Let it be": 81 min)

Danach "verschwand" er aber aus den Regalen, es wurde immer wieder über eine Neuveröffentlichung diskutiert, Bootlegs und Ausschnitte fanden den Weg in die Öffentlichkeit bis schlussendlich 2024 eine Neufassung tatsächlich veröffentlicht wurde und gegenwärtig in Deutschland auf Disney+ zu sehen ist. Verantwortlich für die Neufassung ist Peter Jackson, der zunächst das gesamte Material sichtete und die vielbeachtete Dokumentation "Get back" (siehe nächste Seite) veröffentlichte.

#### Quellen:

https://www.thebeatles.com/let-it-be

https://de.wikipedia.org/wiki/Let\_It\_Be\_(Film)

https://webgrafikk.com/blog/documentary/the-rise-and-fall-and-rise-of-the-let-it-be-film/



Bildquelle (Screenshot):

The Beatles: Get Back - A Sneak Peek from Peter Jackson

https://youtu.be/UocEGvQ100E

Auf Grundlage des vielstündigen Materials, das 1969 für "Let it be" (siehe vorgehende Seite) gedreht wurden, konnte Peter Jackson einen umfangreiche Dokumentation erstellen, der die letzte Schaffensphase der Beatles (als Band) begleitet. Den Abschluss bildet das legendäre Rooftop-Konzert auf dem Dach des Apple-Gebäudes in London am 30. Januar 1969 (Apple, die Beatles-Firma, nicht der MacOS-Hesteller – die gab es 1969 noch nicht). Peter Jackson (und sein Team) waren mehr als vier Jahre mit der Überarbeitung und Zusammenstellung der Aufnahmen beschäftigt.

#### Quellen:

https://de.wikipedia.org/wiki/The\_Beatles:\_Get\_Back

https://de.wikipedia.org/wiki/Rooftop\_Concert

2021



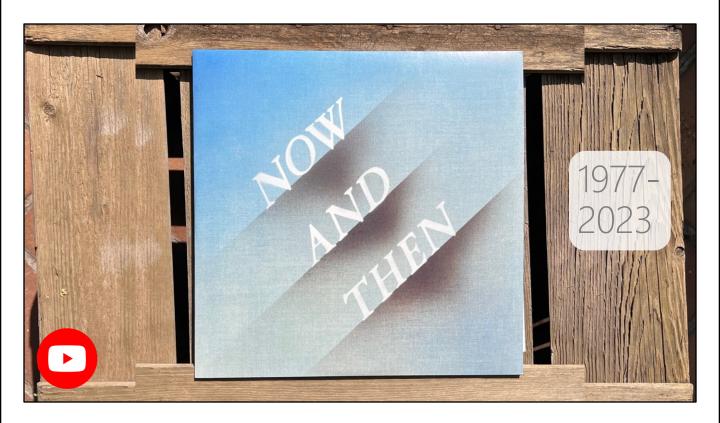
## Bildquelle:

https://www.beatlesmuseum.net/shop/dvds-blu-ray-discs-vhs-videos/blu-ray-discs/beatles-2/the-beatles-blu-ray-get-back/

Die 2021 veröffentlichte Fassung von "Get Back" besteht aus drei umfangeichen Teilen:

- 1. 157 Minuten
- 2. 173 Minuten
- 3. 138 Minuten.

Gesamtlänge der Dokumentation: 468 Minuten



# John Lennons Original Demo:

https://youtu.be/XkLmJ0Gm1hQ

Aufgrund der rechtlichen Situation "verschwinden" viele Videos auf YouTube sehr schnell. Eine Suche nach "John Lennon Now and then original tape recording" mag hilfreich sein.

"Im Januar 1994 übergab Yoko Ono Paul McCartney zwei Musikkassetten mit Heimaufnahmen von John Lennon. Sie enthielten die Lieder Free as a Bird, Real Love, Grow Old With Me und Now and Then. Lennon schrieb Now and Then in den späten 1970er Jahren. Er nahm das unvollendete Musikstück zwischen 1977 und 1979 in Demoform in seiner Wohnung im Dakota Building in New York auf."

Quelle: https://de.wikipedia.org/wiki/Now\_and\_Then\_(Lied)



Rick Beato is Musiker, Musikproduzent und ein sehr bekannter YouTube-Produzent. https://www.youtube.com/@RickBeato

Rick Beator: The Beatles "Now And Then" My Reaction & Breakdown https://youtu.be/7PAkVIFUZPQ

Die hier verwendeten Aufnahmen stammen aus dem oben genannten Video. Ich kann sehr empfehlen, das Video in Gänze zu schauen.

Des Weiteren findet sich auch ein Interview mit Alan Parsons auf Rick Beatos Kanal: Alan Parsons: From The Dark Side of the Moon to The Alan Parsons Project https://youtu.be/xrS6LWHz-dk

Alan Parsons war als Tonassistent bei dem Rooftop-Konzert der Beatles im Einsatz.



Bildquelle: https://youtu.be/7PAkVIFUZPQ

LALAL.AI ist ein Produkt der Schweizer Firma Omnisale GmbH.

"LALAL.AI begann als 2-Stem-Splitter und ist im Laufe des Jahres 2021 erheblich gewachsen.

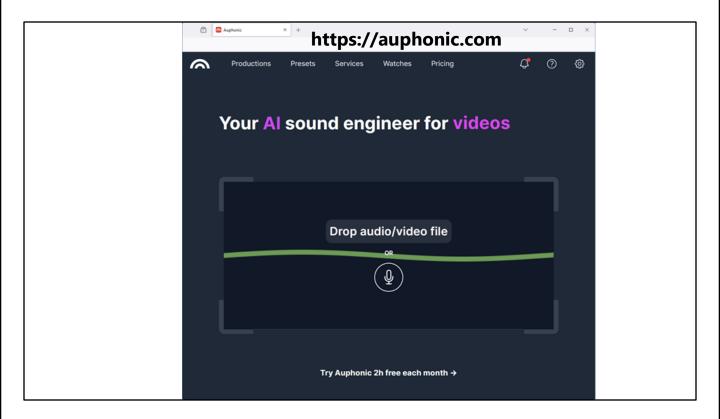
Zusätzlich zu Gesang und Instrumental wurde der Service um die Fähigkeit erweitert,

Musikinstrumente zu extrahieren – Drums, Bass, Akustikgitarre, E-Gitarre, Klavier und Synthesizer.

Infolge dieses Upgrades wurde LALAL.AI zum ersten 8-Stem-Splitter der Welt. Im selben Jahr haben wir auch Geschäftslösungen vorgestellt, die es den Betreibern von Websites, Diensten und

Anwendungen ermöglichen, unsere Technologie der Stem-Aufteilung über eine API in ihre Umgebung zu integrieren."

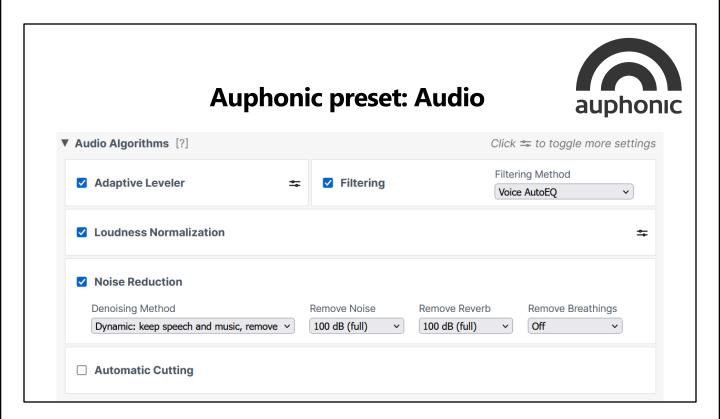
Quelle: https://www.lalal.ai/de/about/



**Auphonic** ist ein Webservice, der im klassischen "Freemium"-Modell die Verbesserung von Tonaufnahmen anbietet. Auphonic ist im Bereich des Podcasting besonders popular. Auphonic integriert zahlreiche Dienste (Whisper, OneDrive, ...) und erlaubt die automatisierte Bereitstellung klangverbesserten Audio- und Videomaterials auf Plattformen wie YouTube.

#### TIPP:

Zwei Stunden pro Monat kann der Dienst kostenfrei genutzt warden.



Zur Analyse und Verbesserung des Tons nutzt Auphonic in vielen Bereich künstliche Intelligenz. Die zahlreichen Algorithmen zur Verbesserung der Audioqualität sind auf den Webseiten des Anbieters ausführlich beschrieben. Eine hervorragende Übersicht bietet diese Seite: https://auphonic.com/features

# Demo

- 1. Lavaliermikro
- 2. Notebookmikro
- 3. <u>Lavaliermikro (Auphonic)</u>
- 4. Notebookmikro (Auphonic)



#### **Broadcast yourself**

Die Testaufnahmen stamen aus einem Vortrag während des MCT Summits 2024 in den Niederlanden. Mein Kollege Ralph Eckhard hat mit mir eine Reihe von Mikrofonen getestet.

Ralph Eckhard: "Calling all admins: Teams Telephony Deep-Dive"

https://www.mctsummit.nl/sessions?sessionId=520090

Ich habe bei diesem MCT Summit sehr ausführlich über die Möglichkeiten gesprochen, Bild- und Tonaufnahmen bei Veranstaltungen mitzuschneiden. Hierzu gibt es nicht nur ein Video namens "Broadcast yourself" sondern auch ein sehr detalliertes Folienset mit vielen Links und Produktbeschreibungen. Alle Verweise finden sich auf meiner Homepage:

http://www.thorsten-butz.de/mctsummit2024

#### Download der Demodateien

Die besprochenen Demodateien finden sich "unbehandelt" unter dem nachfoldenden Link zum Download. Alle Dateien sind ca 1 Minute lang. In dem Vortrag auf der CIM Lingen habe ich die ursprünglichen Demos 2 ("CheapMic") und 3 ("NotebookSound") verwendet:

https://thorsten-butz.de/public/byourself/



#### Windows Server Container - Thorsten Butz - CIM Lingen 2017

https://youtu.be/bFpzrlqhrHg

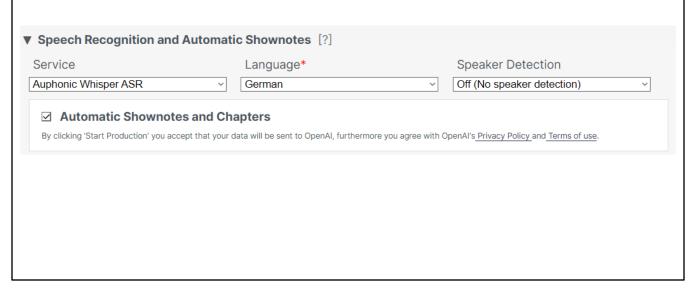
Vortrag auf der CIM Lingen 2017

"Von Docker bis zu Hyper-V Containern: seit Microsoft mit Windows Server 2016 und Windows 10 Container eingeführt hat, sind diese in aller Munde. Wir werfen einen praktischen Blick auf die junge Technologie - nicht nur in Hinblick auf Microsofts Implementierung, sondern auch in Hinblick auf das Original unter Linux.

Dabei gibt es viel Überraschendes zu entdecken: durch die intensive Zusammenarbeit mit der Firma Docker gestaltet sich die Verwaltung der Container für den Windows-Admin denkbar unorthodox: über docker.exe, dem zentralen Kommandozeilenwerkzeug, lassen sich zwar betriebssysemübergreifend alle Aspekte der Konfiguration steuern – die Suche nach PowerShell Cmdlets oder gar einer graphischen Verwaltungskonsole endet jedoch schnell in Ratlosigkeit. Dennoch nimmt die Containerisierung rasant Fahrt auf: mit der Ankündigung des "LinuxKit" schließt sich die vielleicht letzte große Lücke: das Betreiben nativer Linuxinstanzen unter Windows in Containern. Braucht es da eigentlich noch Hyper-V?"







#### **Whisper Speech Recognition in Auphonic**

"So far, Auphonic users had to choose one of our integrated external service providers (Wit.ai, Google Cloud Speech, Amazon Transcribe, Speechmatics) for speech recognition, so audio files were transferred to an external server, using external computing powers, that users had to pay for in their external accounts.

The new Auphonic Speech Recognition is using Whisper, which was published by OpenAI as an open-source project. Open-source means, the publicly shared GitHub repository contains a complete Whisper package including source code, examples, and research results.

However, automatic speech recognition is a very time and hardware-consuming process, that can be incredibly slow using a standard home computer without special GPUs. So we decided to integrate this service and offer you automatic speech recognition (ASR) by Whisper processed on our own hardware, just like any other Auphonic processing task, giving you quite some benefits:

No external account is needed anymore to run ASR in Auphonic.

Your data doesn't leave our Auphonic servers for ASR processing.

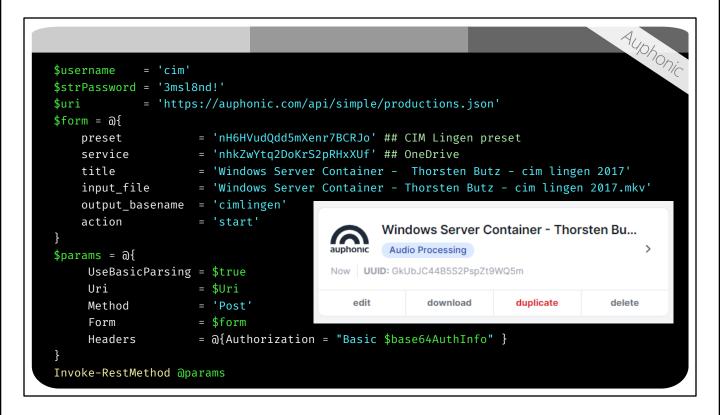
No extra costs for external ASR services.

Additional Auphonic pre- and post-processing for more accurate ASR, especially for Multitrack Productions.

The quality of Whisper ASR is absolutely comparable to the "best" services in our comparison table."

#### Quelle:

https://auphonic.com/blog/2022/11/08/auphonic-whisper-asr-beta/



#### **Auphonic API**

Auphonic bietet eine REST API, mit deren Hilfe man die Verarbeitung der rohen Audio- oder Videodateien automatisieren kann. Der hier beschriebene recht einfache Weg basiert auf drei Schritte:

- Erstellen eines Presets in Auphonic (graphisch), der alle nötigen Parameter enthält, die später auf die Quelldateien angewendet weden soll.
- 2. Upload der Quelldateien in einen Cloudspeicher, zum Beispiel OneDrive.
- 3. Ausführen des PowerShell-Snippets.

https://auphonic.com/help/api/simple\_api.html#start-a-production-and-upload-a-file

Natürlich mag man für einzelne Dateien den Weg über die graphische Oberfläche vorziehen. Auf die gezeigte Weise ist es aber auch möglich, in schneller Folge beliebig viele Videos zu Auphonic zu senden und von dort aus zu publizieren, zum Beispiel nach YouTube.

unbearbeirer

```
WEBVTT
```

00:00:00.000 --> 00:00:06.960 Music.

00:00:06.808 --> 00:00:12.588

Thorsten, ganz herzlich willkommen hier auf der True Metal Stage,

00:00:12.768 --> 00:00:13.688 heißt das glaube ich, oder?

00:00:13.748 --> 00:00:17.668

Oder in der Sprache der **Zimling**, der Tech-Track.

00:00:19.368 --> 00:00:25.748

Mein Name ist Thorsten Butz. Ich spreche über ein sicherlich sehr besonderes Thema.

00:00:25.748 --> 00:00:31.488

Gut, das tun wir alle, aber wir verbinden in dieser Session zwei Welten,

00:00:31.488 --> 00:00:36.948

die so in meiner Adoleszenz als IT-Mensch immer sehr getrennt waren.

Auphonic generiert mit Hilfe der Whisper-KI auf Wunsch Untertitel. Alternativ zum Format VTT lassen sich auch SRT-Dateien genierieren. Der gleiche Abschnitt sieht in SRT dann wie folgt aus:

```
1
00:00:00,000 --> 00:00:06,960
Music.
2
00:00:06,808 --> 00:00:12,588
Thorsten, ganz herzlich willkommen hier auf der True Metal Stage,
3
00:00:12,768 --> 00:00:13,688
heißt das glaube ich, oder?
00:00:13,748 --> 00:00:17,668
Oder in der Sprache der Zimling, der Tech-Track.
00:00:19,368 --> 00:00:25,748
Mein Name ist Thorsten Butz. Ich spreche über ein sicherlich sehr besonderes Thema.
6
00:00:25,748 --> 00:00:31,488
Gut, das tun wir alle, aber wir verbinden in dieser Session zwei Welten,
7
00:00:31,488 --> 00:00:36,948
die so in meiner Adoleszenz als IT-Mensch immer sehr getrennt waren.
```

Zimling Change-Route Linux-Körnel GetProcess Drüppel Kettles & Pads

CIM Lingen chroot Linux kernel Get-Process Drupal Cattles & pets

#### Fehler der Transkription

Die Fehlerquote ist überraschend gering, aber natürlich findet man im Rohmaterial noch eine ganze Reihe von Problemfällen. Es ist nachvollziehbar, dass die Mischung aus Deutsch und Englisch, aus Code und "normaler" Sprache eine maximale Herausforderung für die Transkription ist. Oben sieht man einige Fehler aus der Rohfassung.



Title (required)

Windows Server Container - Thorsten Butz - cim lingen 2017 (remastered)

#### Description

In dieser Vorlesung spricht Thorsten Butz über die Schnittstelle zwischen klassischer IT und modernem Container-Management, insbesondere in Bezug auf Windows Server Container und Docker. Er erläutert, wie in seiner Jugend die IT-Welt durch unterschiedliche Systeme wie Linux und Windows getrennt war, und wie Container-Technologien diese Grenzen aufbrechen. Der Vortrag beginnt mit einer Erklärung, dass es nicht nur um Windows Server Container geht, sondern auch um Docker und die zugrunde liegenden Technologien.

#### KI-generierte Zusammenfassungen

Auphonic/Whisper generieren neben Untertiteln auch Zusammenfassungen des Inhalts, was die Nützlichkeit der Spracherkennung auf eine völlig neue Ebene bringt. In YouTube finden sich die Zusammenfassungen dann direct in der Beschreibung einschließlich (wenn gewünscht) automatisch generierten Kapitelmarken.

#### **Demovideo auf YT** (nur direkt über den Link aufrufbar)

Windows Server Container - Thorsten Butz - cim lingen 2017 (remastered)

https://youtu.be/egtrQg5tuF0

Der gleiche Inhalt findet sich auf meiner Homepage noch einmal, dort verlinke ich das Video mit der werbefreien Plattform Vimeo.

### Kapitelmarken (Ausschnitt):

00:00:06.309 Willkommen auf der True Metal Stage

00:06:48.912 Grundprinzip von Containern

00:09:06.292 Virtualisierung und deren Entwicklung

```
$API_KEY = 'just-an-example:key'
$uri = 'https://api-free.deepl.com/v2/translate'

$text = @'
Hallo Lingen!
Hallo Welt!
'@

$headers = @{
   'Authorization' = "DeepL-Auth-Key $API_KEY"
}

$body = @{
        'text'=$text
        'target_lang'='EN'
}

Invoke-RestMethod -Method 'Post' -Headers $headers -Uri $uri -Body $body
```

#### Übersetzungen

Der deutsche Anbieter DeepL bietet seit 2017 einen KI-gestützten Übersetzungsdienst, der durch herausragende Qualität besticht. Auch DeepL bietet einen kostenfreien Basiszugang und eine REST-API zur Automation.

Die Codebeispiele finden sich als Textdateien im Repo.

#### Links und Co:

https://github.com/thorstenbutz/conferences/tree/master/2024.CIMLingen https://thorsten-butz.de



Windows Server Container - Thorsten Butz - CIM Lingen 2017

Vortrag auf der CIM Lingen 2017

## Ursprüngliche Version aus 2017

https://youtu.be/bFpzrlqhrHg

#### Remastered von 2024

https://youtu.be/egtrQg5tuF0

https://vimeo.com/thorstenbutz/cimlingen2017

Die Neuauflage ("remastered") wurde mit Hilfe von **Auphonic** klanglich optimiert, des Weiteren wurde ein Transkript, Untertitel und eine Zusammenfassung erstellt.

Für die veröffentlichte Version habe ich händisch einige Transkriptionsfehler beseitigt. Auf Grundlage der korrigierten deutschen Untertiteldatei habe ich mit Hilfe von **DeepL** eine englische Übersetzung erstellen lassen (ohne weitere Korrekturen) und als zusätzliche Untertitelspur hinzugefügt.



#### **PSConfEU** @psconfeu.bsky.social · 2mo

99 videos are available - not too bad, isn't it?

htp./ponershellvideo But we're still not completely done. We have a couple of interviews in our queue. From Walter to Steven, from Amber to Bruce.

And of course: Voices from the community aka from YOU!

We will publish 1 more video every day!

#PSConfEU



#### http://powershell.video

https://youtube.com/playlist?list=PLDCEho7foSor3zw2nZVi4AR100bRr6Ehs

Die europäische PowerShell-Konferenz (psconf.eu) ist seit 2016 das Stelldichein der Szene. Mit typischerweise 4 parallelen Tracks an 4 Konferenztagen entstehen eine Großzahl an Aufnahmen, die kostenfrei auf dem YT-Kanal angeboten werden.

Ich bin seit einigen Jahren der Verantwortliche für das Bereitstellen dieser Aufnahme. Im Jahr 2024 sind in Antwerpen ca 100 Videos entstanden. Neben den zahlreichen Vorträgen sind auch eine Reihe von Interviews und "Behind the scenes .."-Aufnahmen entstanden, die von Hand geschnitten werden.

Den Großteil der Vorträge bearbeite ich aber ausschließlich mit der Kommandozeile. Mit Hilfe von PowerShell, FFmpeg, Auphonic und Co kann ich Videos in großer Zahl veröffentlichen. In 2024 waren mehr als die Hälfte der Videos wenige Tage nach der Konferenz online, inclusive (KI-generierter) Zusammenfassungen/Beschreibung und optimiertem Ton.



#### **Abbey Road**

Das Bild zeigt mich 2023 an der Abbey Road, auf dem berühmten "Crosswalk", der das vorletzte Beatlesalbum zierte. In der Bildmitte links (das weiße Gebäude, vor dem die vielen Menschen stehen) ist das berühmte Tonstudio, in dem die Beatles zahlreiche Alben aufgenommen haben – unter anderem "With the Beatles" im Jahr 1963. Seinerzeit hieß das Studio noch: "EMI Recording Studios"

#### Weitere Vorträge zur Audio-/Videoaufzeichnung und -bearbeitung

- about\_VideoProcessing Thorsten Butz PSConfEU 2022
   https://youtu.be/i4watr8txOM
- about\_VideoProcessing, Season 2 Thorsten Butz PSConfEU 2023
   https://youtu.be/B8ubWEyYbXM
- Broadcast yourself Thorsten Butz MCT Summit 2024
   https://vimeo.com/thorstenbutz/byourself
   https://www.thorsten-butz.de/mctsummit2024/

# wir danken unseren cim sponsoren!



















#### Binge-watching The Beatles

- Marti Fischer: The Beatles Now And Then (1970 Version) https://youtu.be/CKkd9BCjz68
- The Beatles Now And Then (Official Music Video) https://youtu.be/Opxhh9Oh3rg
- John Lennons Original Demo "Now and then": https://youtu.be/XkLmJ0Gm1hQ
- The Beatles All My Loving (Live at the Morecambe and Wise Show, 1963) https://youtu.be/6N4frdY1h2w
- The Beatles: Get Back A Sneak Peek from Peter Jackson https://youtu.be/UocEGvQ100E
- Rick Beator: The Beatles "Now And Then" My Reaction & Breakdown https://youtu.be/7PAkVIFUZPQ
- Get back (Official Trailer) https://youtu.be/Auta2lagtw4
- Let It Be (Official Trailer) https://youtu.be/FvR2 CCoX60

# Extra-Tipp:

James Corden, Paul McCartney
 Carpool Karaoke