IDEE & ENTSTEHUNG

# Nicht nur kürzen. Eine Geschichte erhalten.

Wie mein KI-gestützter Longform-Workflow funktioniert

Aus langen RP-Streams sollen verständliche Folgen werden: mit Figuren, Humor, Atmosphäre und einem natürlichen Ende. Nicht einfach die lautesten Momente hintereinander - und auch nicht der komplette Stream mit einem neuen Dateinamen.

## Die Grundidee

Ich verbinde einen KI-Assistenten in Hermes Agent mit DaVinci Resolve Studio und spezialisierten Medienwerkzeugen. Der Assistent hilft bei Sichtung, Episodenplanung, Schnittumsetzung, Prüfung und Verpackung. Ich gebe die kreative Richtung vor und beurteile, ob das Ergebnis tatsächlich funktioniert.

> **Kein einzelner Wunder-Prompt: Der Workflow ist aus vorhandenen Werkzeugen, konkreten Schnittaufträgen und wiederholtem Feedback entstanden.**

## Wie ich ihn verwirklicht habe

• Das Ergebnis zuerst beschrieben: ruhiges Erzählen, verständliche Zusammenhänge, vollständige Reaktionen und Respekt vor den Mitspielenden - statt möglichst vieler Schnitte.

• Vorhandenes verbunden: Aufnahmen, Transkripte, Markenassets und entwickelte Folgen wurden zur Grundlage. Der funktionierende Hermes-/Resolve-Weg musste nicht auf die Fertigstellung eines großen eigenen Gesamtsystems warten.

• Die technische Umsetzung mit KI erarbeitet: Der Assistent hat Werkzeuge genutzt und ergänzende Skripte für Materialordnung, Zeitbereiche, Schnittaufbau und Prüfung erstellt. Kleine Programme übernehmen wiederholbare Arbeit; das Sprachmodell hilft bei der Einordnung.

• Am tatsächlichen Film korrigiert: Ein unverständlicher Einstieg, zu wenig hörbare Atmosphäre oder ein falscher Schluss führten zu konkreten Änderungen. Bewährte Entscheidungen wurden in Projektunterlagen und wiederverwendbaren Arbeitsanweisungen festgehalten.

## Wer übernimmt was?

| Ich als Creator | Assistent und Werkzeuge |
|---|---|
| Absicht, Ton und kreative Grenzen | Material erschließen und Auswahl begründen |
| Feedback am wirklichen Bild und Ton | Schnitt umsetzen, Exporte und Prüfnachweise liefern |
| Kreative Endbeurteilung und Veröffentlichung | Projekt, Dateien und Veröffentlichungstexte vorbereiten |

Der entscheidende Wechsel: Ich erkläre nicht jeden einzelnen Mausklick. Ich beschreibe das gewünschte Ergebnis und korrigiere am konkreten Beispiel. Das ersetzt nicht meinen Geschmack - es macht ihn zum Maßstab der technischen Arbeit.

ABLAUF & WERKZEUGE

# Vom Rohmaterial zum fertigen Paket

> **Aufnahme → Transkript → Episodenplan → Resolve → Prüfung → Paket**

## 1 · Das Material eindeutig zuordnen

Aufnahmen werden nach Quelle und zeitlichem Zusammenhang geordnet. Doppelte Containerdateien sind nicht automatisch zwei verschiedene Streams. Originale bleiben erhalten; Arbeitsstände und Exporte liegen getrennt davon.

## 2 · Sprache durchsuchbar machen

Vorhandene Transkripte werden wiederverwendet; fehlende oder unklare Stellen lassen sich lokal mit Whisper beziehungsweise faster-whisper ergänzen. Zeitmarken helfen, Gespräche wiederzufinden. Ein Transkript ist ein Suchindex - keine fehlerfreie Mitschrift und keine vollständige Videosichtung.

## 3 · Eine Folge redaktionell auswählen

Der Assistent sucht zusammenhängende Handlungs- und Gesprächsbögen: Was beginnt hier? Was brauchen Zuschauende zum Verständnis? Wo endet ein vollständiger Moment? Die Auswahl hält fest, welche Originalbereiche bleiben, was entfällt und welche Fragen noch offen sind.

## 4 · Knackpunkte gezielt ansehen und anhören

Einstiege, Übergänge und Schlussstellen werden mit Vor- und Nachlauf geprüft. Lokale Bilder, Tonanalysen und bei Bedarf kurze audiovisuelle Modellprüfungen helfen dabei. Gemini dient für konkrete Detailfragen an kurzen Ausschnitten - nicht zum pauschalen Hochladen stundenlanger VODs.

## 5 · In Resolve umsetzen und gestalten

DaVinci Resolve Studio ist das eigentliche Schnittsystem. Über die angebundene Skript-/Werkzeugschnittstelle, unter anderem MCP, kann der Assistent native Timelines aufbauen und prüfen. MCP ist dabei die Verbindung zu Werkzeugen, keine zusätzliche Schnitt-KI. Nur eine Instanz schreibt an der aktiven Timeline. Bestehender Look, Markenkarten und passende Musik werden übernommen oder gezielt ergänzt. Über die Epidemic-Sound-Anbindung sucht und lädt der Assistent Musik und Effekte und importiert sie anschließend in Resolve; entscheidend ist deren hörbare Funktion, nicht die Anzahl belegter Tonspuren. Viele Serienfolgen nutzen bewusst nur kurze Kartenmusik und lassen Dialoge frei.

## 6 · Den wirklichen Export prüfen

FFmpeg und ffprobe prüfen unter anderem Lesbarkeit, Laufzeit, Bildrate, Ton und auffällige Schwarz-, Stille- oder Stillstandsbereiche. Native Timeline-Daten werden mit dem Plan verglichen. Auffälligkeiten sind Hinweise zur Prüfung, nicht automatisch Fehler.

## 7 · Mehr liefern als eine MP4

Zum Paket gehören Master, kompakte Ansicht, editierbares Projekt beziehungsweise Timeline, Thumbnail, Titel-/Beschreibungsvorschlag, Kapitel und ein kurzer Prüfstand. Ein Resolve-Projektexport allein enthält die ursprünglichen Medien nicht automatisch.

PRAXIS & QUALITÄT

# Die Schnittentscheidung ist wichtiger als der Effekt

## Ein echtes Beispiel: Folge LF10

„Formation, Pferde und ein schwieriger Befehl“ enthält Formations- und Pferdegespräche, eine Depesche mit Eilritt und eine zusammenhängende Diskussion. Die Folge wurde nicht auf eine starre Ziellänge zusammengestrichen.

| Auswahl aus der Originalaufnahme | Entscheidung |
|---|---|
| Sekunde 3782-4371 und 4444-6433 | Zwei zusammenhängende Originalblöcke behalten |
| Sekunde 4371-4444 | 73 Sekunden sprachfreie Wartephase entfernt |
| 42:58 Minuten beibehaltenes Originalmaterial | Gespräch mit Gegenrede und Nachreaktionen erhalten |
| 43:11 Minuten fertige Schnittfassung | Mit 5 Sekunden Titel und 8 Sekunden Endkarte |

Das Beispiel zeigt: Der Workflow ist nicht nur eine automatische Highlight-Maschine. Ein längerer vollständiger Bogen kann die bessere Folge sein. Umgekehrt werden kurze Geschichten nicht künstlich auf eine Mindestlänge aufgefüllt.

## Vier Regeln, die im Alltag den Unterschied machen

• Ursache vor Reaktion: Wenn ein Dialog auf eine Explosion reagiert, muss der ursprüngliche Knall im Einstieg auch hörbar sein. Nur die Reaktion zu behalten, zerstört den Zusammenhang. Ein beliebiger nachträglicher Effekt ist kein Ersatz für das fehlende Originalereignis.

• Vollständige Sätze und Reaktionen: Wortzeitmarken helfen beim Suchen, treffen aber keine garantierten Silbengrenzen. Vorlauf, Atem, Wortende und kurze Nachreaktionen müssen geschützt werden.

• Figur und Privatperson unterscheiden: IC bedeutet innerhalb der gespielten Rolle, OOC außerhalb davon. Ein Spieler-Nachgespräch darf nicht mit warmer Musik als Entwicklung der Figur verkauft werden. Unsichere Mischgespräche werden nicht durch einzelne herausgelöste Antworten passend gemacht.

• Atmosphäre muss ankommen: Eine vorhandene Musikdatei beweist keine Stimmung. Manchmal helfen wenige gezielte Akzente; manchmal ist Originalton die stärkere Gestaltung. Bei strittigen Stellen wird ein kurzer echter Bild-und-Ton-Ausschnitt beurteilt, bevor erneut die ganze Folge rendert.

> **Render erfolgreich ≠ Folge kreativ gelungen. Technische Prüfung und kreative Endbeurteilung beantworten unterschiedliche Fragen.**

## Was die Prüfung leisten kann - und was nicht

Ein vollständiger Decode kann beschädigte Videodaten aufdecken. Lautheitsmessung zeigt Pegel und Reserven. Ein Schwarzbilddetektor findet Kandidaten für Lücken. Aber eine Spielkarte kann absichtlich stillstehen, eine Szene dunkel sein und eine Pause Bedeutung haben. Der Kontext entscheidet. Die Produktionsberichte trennen deshalb technische Tests, visuelle Stichproben und begrenzte audiovisuelle Detailprüfungen von einer vollständigen menschlichen Abnahme. Auch die fertige LF10-Schnittfassung war damit nicht automatisch veröffentlicht oder vollständig kreativ freigegeben.

SELBST AUSPROBIEREN

# Klein anfangen. Am Ergebnis lernen.

## Was man dafür braucht

• Eine gute Aufnahme: Verständlicher Originalton ist wichtiger als ein späterer KI-Reparaturwunsch. Dazu ausreichend Speicher für Originale, Arbeitsdateien, Exporte und Sicherungen.

• Einen handlungsfähigen Assistenten: Hier Hermes Agent mit geeignetem Sprachmodell und freigegebenem Werkzeugzugriff. Ein normaler Textchat ohne Datei- und Schnittzugriff erledigt diese Arbeit nicht allein.

• Eine funktionierende Werkzeugkette: Für die hier gezeigte Anbindung verwende ich die kostenpflichtige Studio-Ausgabe von DaVinci Resolve, ergänzt um Python, FFmpeg und ffprobe.

• Transkripte nach Bedarf: Vorhandene Transkripte lassen sich wiederverwenden. Lokale Transkription kommt dazu, wenn sie gebraucht wird. Musik und Markenassets richten sich nach der Gestaltung.

Mein Studio-PC ist die Referenz, keine allgemeine Mindestanforderung. Leistung beeinflusst Wartezeiten, mögliche Modelle und Parallelität. Ein zusätzlicher Server oder ein komplettes eigenes Creator-Betriebssystem ist für das Grundprinzip keine Voraussetzung.

## Ein sinnvoller erster Versuch

Eine überschaubare Szene aus einer Aufnahme wählen, vorhandenen Kontext mitgeben und zunächst genau eine Folge oder einen kurzen Piloten herstellen. Erst wenn Auswahl, Übergänge, Ton und Export funktionieren, den Ablauf auf mehr Material ausweiten.

> **Beispielauftrag, keine wörtliche historische Eingabe: „Erstelle aus dieser Aufnahme eine eigenständig verständliche Folge. Erhalte Sprechtempo, Ursachen und vollständige Reaktionen. Keine erfundenen Zusammenhänge. Lege zuerst Quellbereiche und offene Fragen fest, prüfe kritische Stellen mit Vor- und Nachlauf und liefere eine editierbare Fassung plus Ansicht. Nicht veröffentlichen.“**

## Drei Dinge vorher klären

• Kosten: Modellnutzung, audiovisuelle Cloudprüfung, Resolve-Studio-Lizenz und Musiklizenz können Geld kosten. Für diesen Workflow ist hier keine belastbare Pauschale pro Folge angegeben. Kleine Testausschnitte und die Wiederverwendung vorhandener Analysen begrenzen unnötigen Aufwand.

• Daten und Rechte: Lokaler Schnitt bedeutet nicht automatisch rein lokale KI-Verarbeitung. Je nach Modellweg können Text und ausgewählte Ausschnitte an Cloudanbieter gehen. Vorher festlegen, was wohin darf; Rechte an Musik, Spielmaterial und Beiträgen anderer getrennt prüfen.

• Kontrolle: Originale sichern, Änderungen versionieren und Werkzeuge nur im nötigen Umfang freigeben. Automatische Herstellung ist keine automatische Erlaubnis zum Hochladen oder Veröffentlichen.

## Der übertragbare Kern

Nicht mein Rechner und nicht ein bestimmtes Modell sind die eigentliche Methode. Es ist die Verbindung aus klarer kreativer Absicht, ausführbaren Werkzeugen, wiederverwendbaren Regeln und überprüften Ergebnissen. Andere können dieses Prinzip übernehmen - ihre Anbindung und Ergebnisqualität müssen sie trotzdem selbst erproben.

Die konkrete Komponentenübersicht und anklickbare Projektlinks findest du auf Seite 5. Dieses Handout beschreibt einen individuell entwickelten Praxisworkflow, kein fertiges Ein-Klick-Produkt und keine universelle Installationsanleitung.

KOMPONENTEN & LINKS

# Mein Aufbau im gezeigten Video

Die verwendeten Werkzeuge und ihre Aufgaben. Versionsangaben beziehen sich auf die gezeigte Produktion.

## Grundaufbau

| Komponente | Verwendung in meinem Aufbau |
|---|---|
| Hermes Agent | Umgebung für Dateien, Arbeitsanweisungen und Werkzeugaufrufe. |
| gpt-6.1-sol | Hauptmodell über meinen OpenAI-Abozugang: Planung, Transkriptanalyse, Skripte und Werkzeugaufrufe. Andere geeignete Modelle sind möglich. |
| DaVinci Resolve Studio 21.1 | Kostenpflichtiges Schnittprogramm für die gezeigte Produktion. Die Timeline bleibt editierbar. |
| Resolve MCP | In meinem Aufbau verwendete MCP-Anbindung: Skripte ausführen, Medien importieren, Timelines bearbeiten und Renderjobs starten. |
| Python · FFmpeg · ffprobe | Skripte und Schnittdaten verbinden, Prüfausschnitte und Ansichten erstellen, Audio bearbeiten und Dateieigenschaften auslesen. |

## Ergänzungen nach Bedarf

| Komponente | Aufgabe und Einordnung |
|---|---|
| Whisper / faster-whisper | Lokale Transkription. Für gezielte Nachprüfungen nutzte ich large-v3 auf der GPU; ergänzend einen kleineren Lauf auf der CPU. Vorhandene Transkripte können ausreichen. |
| Gemini über Gemini Video | Kurze audiovisuelle Detailprüfungen über eine eigene MCP-Anbindung. Bei den gezeigten Prüfungen: vier Bilder pro Sekunde. Ergänzung für offene Fragen. |
| Epidemic Sound | Musik und Effekte per Anbindung suchen, herunterladen und in Resolve importieren. Im Video: „Backyard Folklore“ von Northside, auch als Instrumental-Stem. Passende Lizenz erforderlich. |
| Qwen über Strata | Zusätzlicher lokaler Modellweg auf meinem Rechner mit einer 5090. Für den Einstieg nicht erforderlich. |
| Three.js | KI-unterstützt programmierte 3D-Erklärszenen für dieses Infovideo, anschließend gerendert und in Resolve montiert. Für den eigentlichen Streamschnitt nicht erforderlich. |

## Weiterführende Projektseiten

Hermes Agent · DaVinci Resolve · Python · FFmpeg / ffprobe

Whisper · faster-whisper · Epidemic Sound · Three.js

Resolve MCP und Gemini Video bezeichnen die Anbindungen meines eigenen Aufbaus. Die Projektlinks ersetzen deren Einrichtung nicht. Interne Pfade und Zugangsdaten sind nicht Bestandteil dieses Handouts.