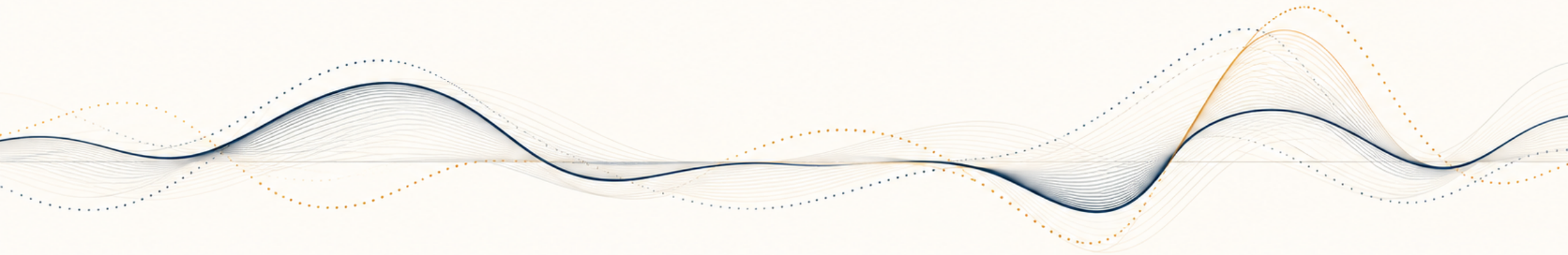


KI hört mit - und mischt sich ein

Was künstliche Intelligenz im Audioworkflow wirklich kann



Broadcast hat enge Regeln

Ein gutes Ergebnis reicht nicht. Es muss immer wiederholbar funktionieren.

guter Klang

Live

viele Kanäle

zuverlässig

vorhersagbar

bezahlbar

**Broadcast-tauglich heißt:
Qualität unter Lastbedingungen**

Die technische Frage ist aber nicht: „Kann KI Audio?“
Sondern: „Kann sie es live, stabil und wirtschaftlich?“

KI, Machine Learning, Deep Learning - nicht ganz dasselbe

KÜNSTLICHE INTELLIGENZ

Überbegriff für Funktionen: Regeln, Suche, Wahrnehmung, Entscheiden, Lernen, hybride Ansätze

MACHINE LEARNING

Modelle lernen Muster aus Daten - statt jede Regel explizit zu programmieren

DEEP LEARNING

Machine Learning mit mehrschichtigen neuronalen Netzen

Generative AI

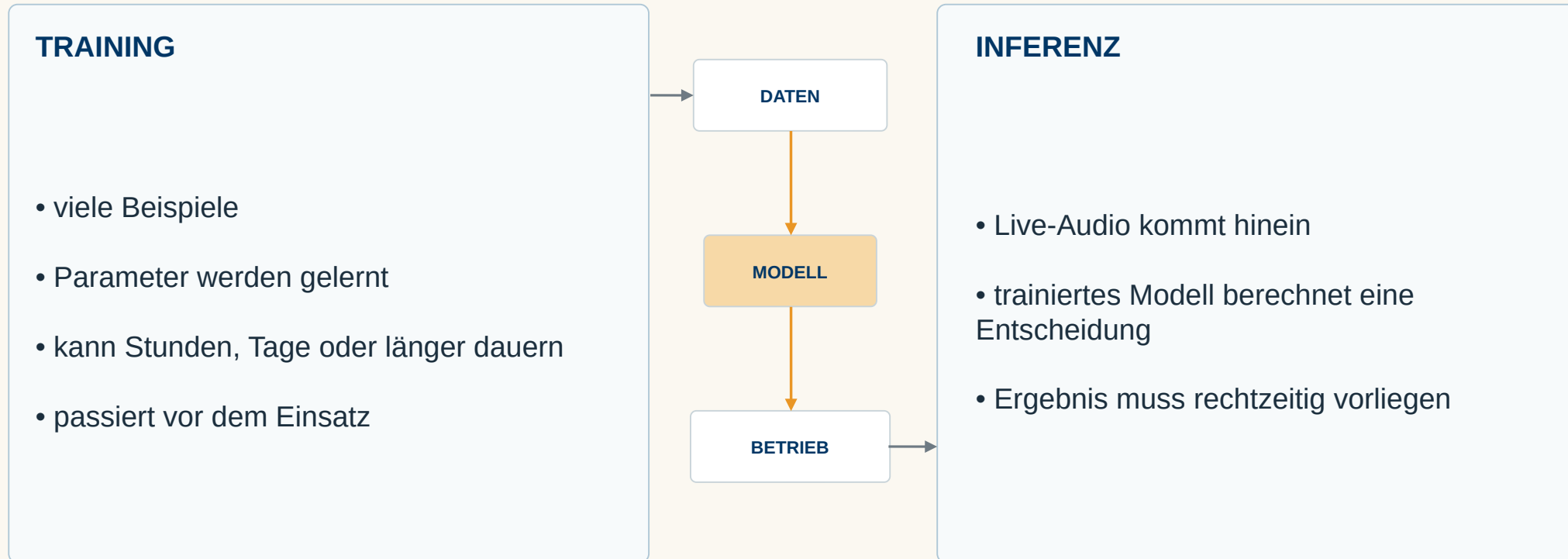
Aufgabenklasse: erzeugt Inhalte, Text, Sprache, Bilder, Musik...

LLMs

Sprachmodelle sind prominent, aber nicht das Grundmodell für Audio-DSP

Wenn im Audio von „KI“ die Rede ist, steckt dahinter meist ML oder Deep Learning

Training und Inferenz



**Zwei unterschiedliche Kostenprobleme:
Trainingsaufwand vs. Betrieb pro Kanal, pro Sekunde, pro System.**

Was ML mit Audio machen kann

1

ERKENNEN

Sprache · Musik · Sprecher · Geräusch

2

BEWERTEN

Qualität · Verständlichkeit · Störung

3

TRENNEN / VERBESSERN

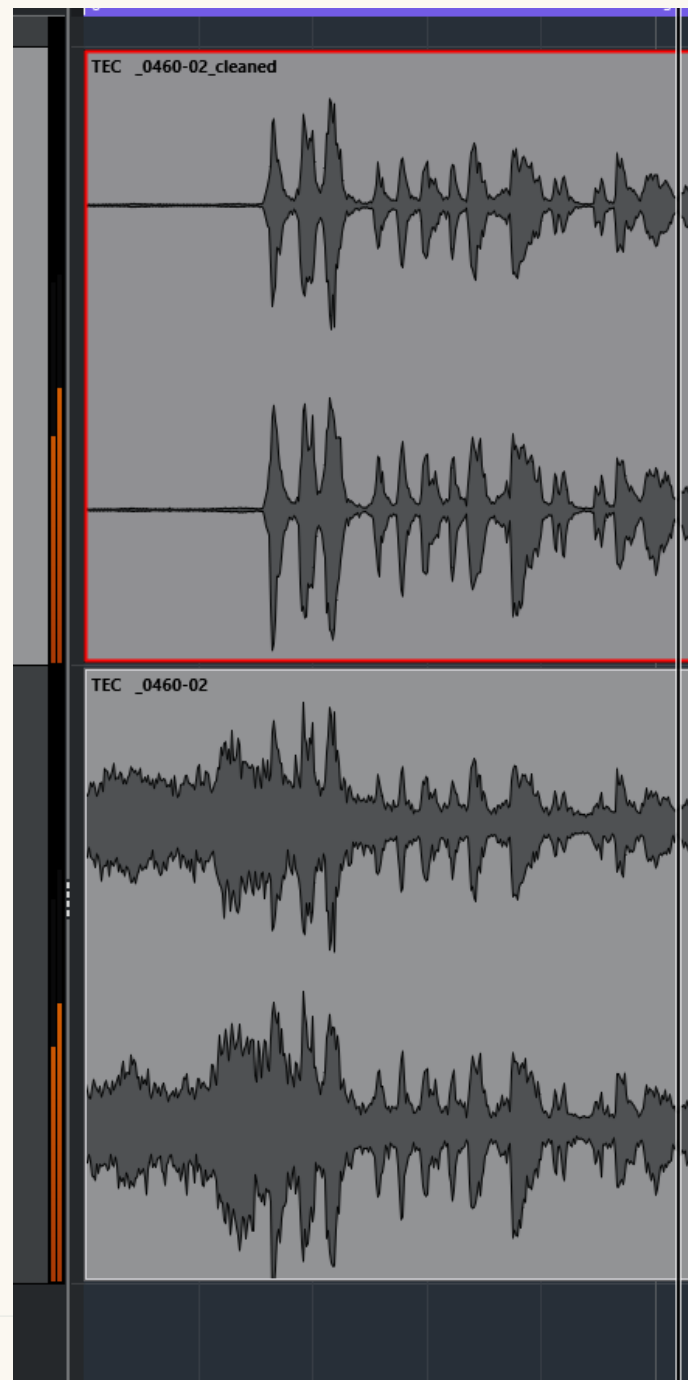
Denoising · Enhancement · Quelltrennung

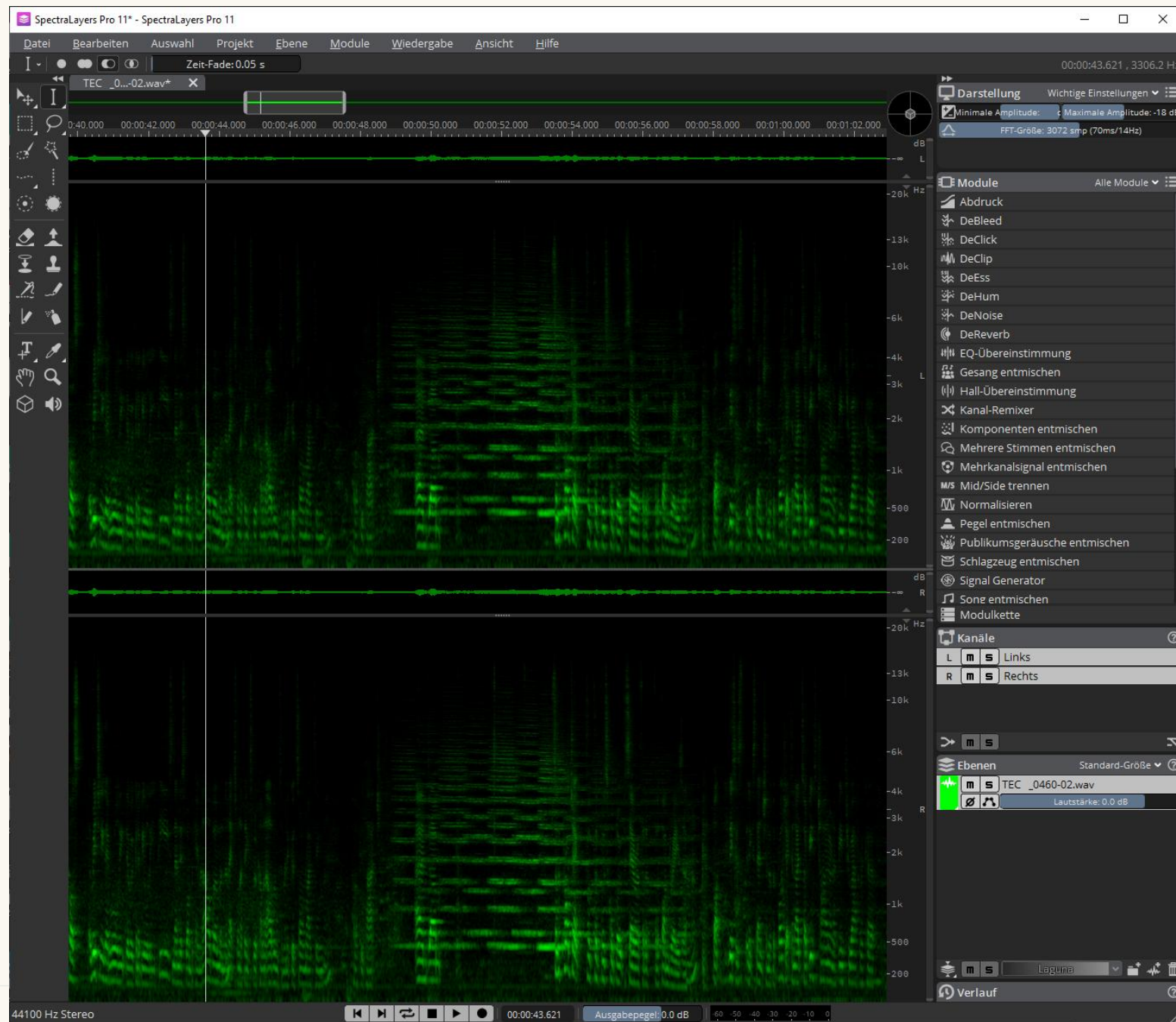
4

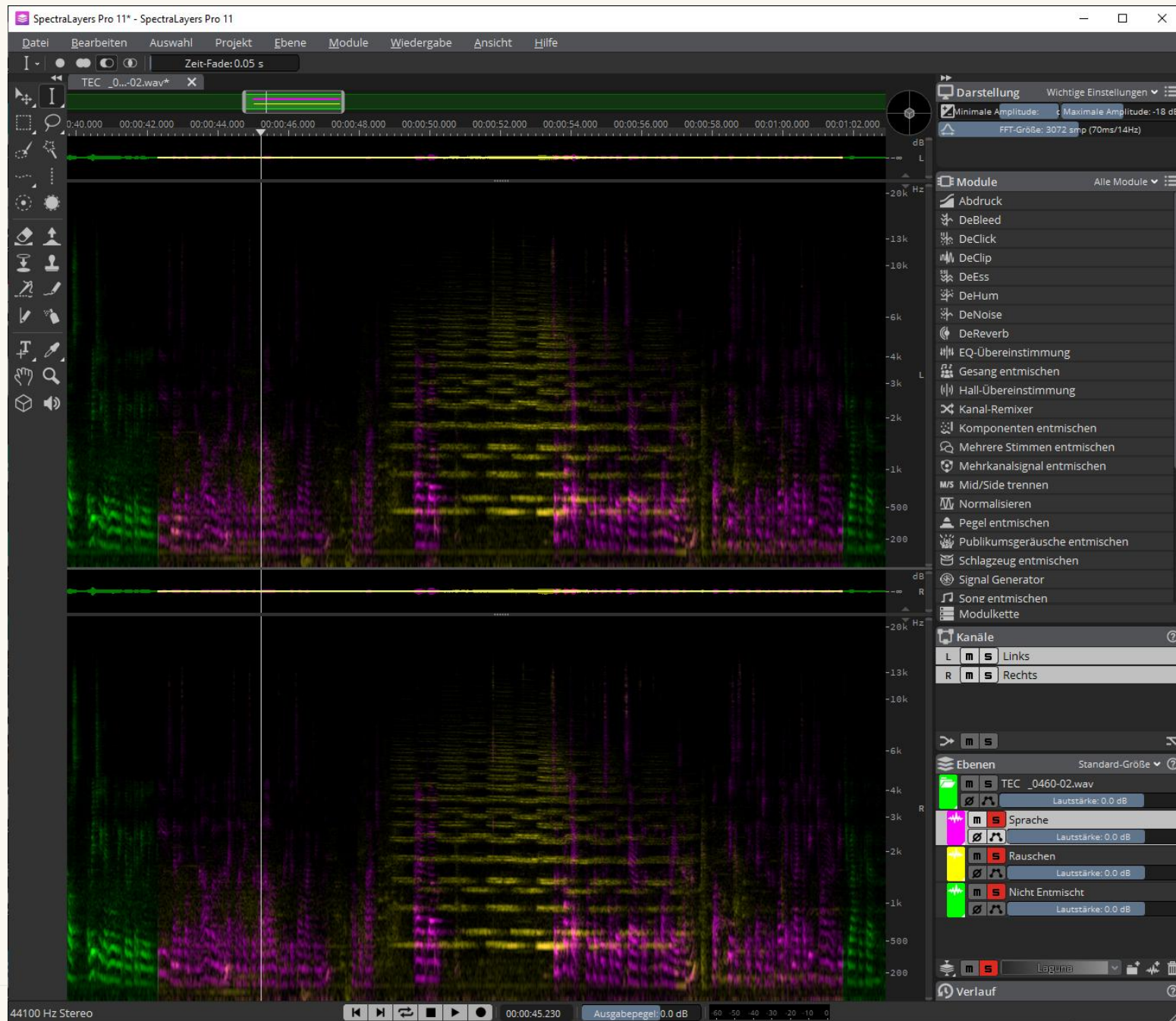
ERZEUGEN

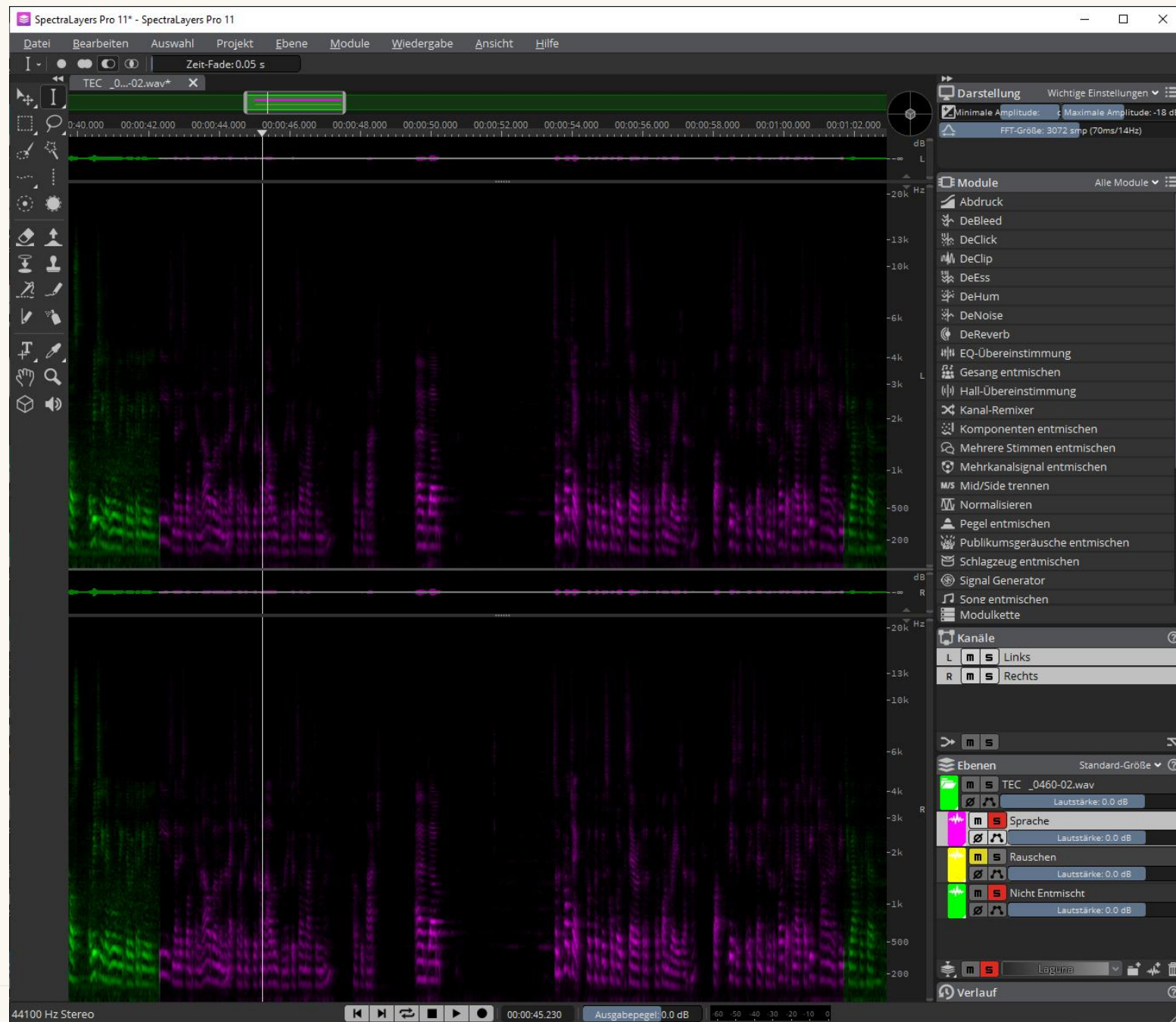
Sprache · Musik · Klang

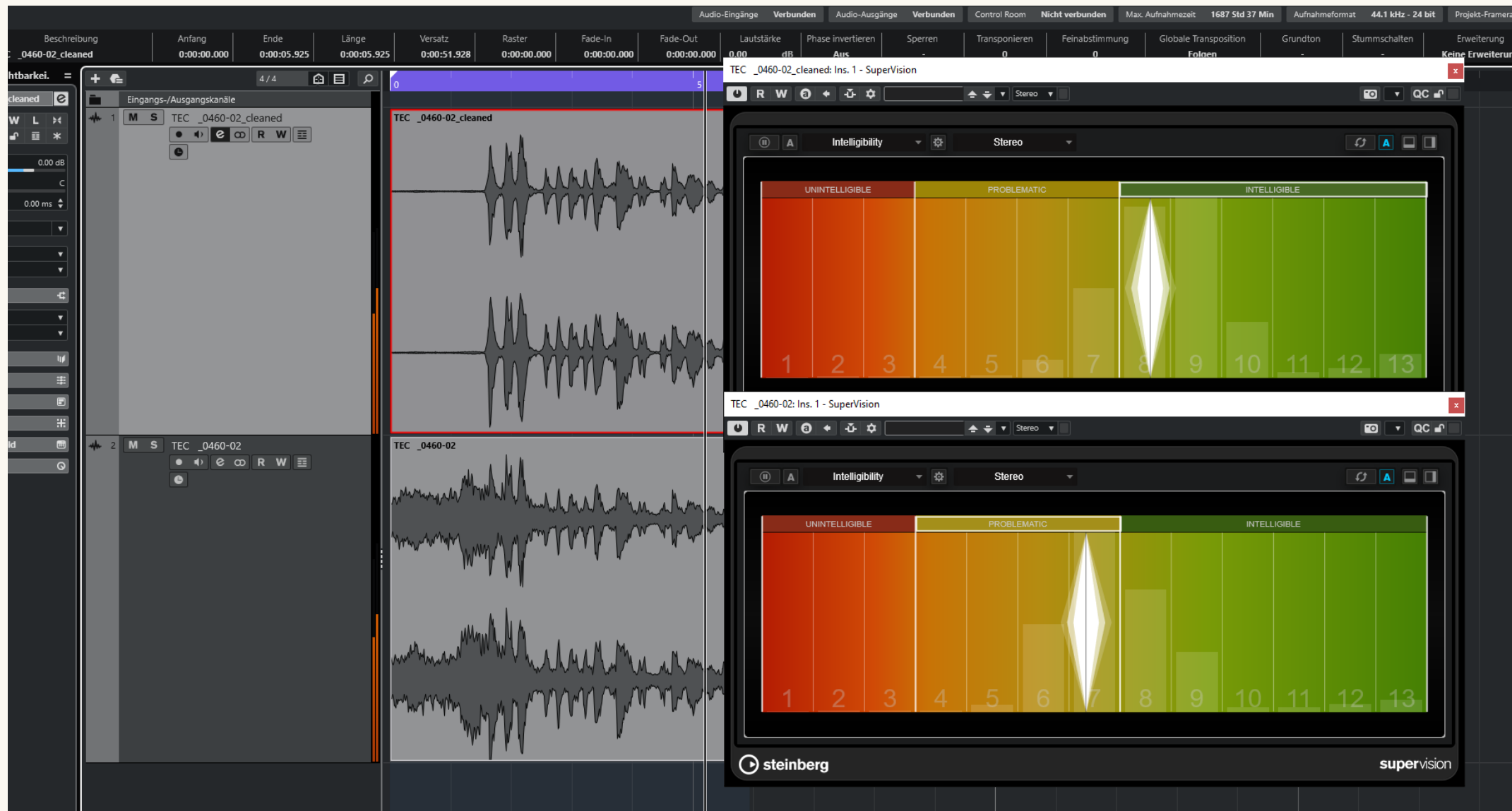
Nicht jede dieser Aufgaben gehört automatisch in den Echtzeit-Audiopfad.

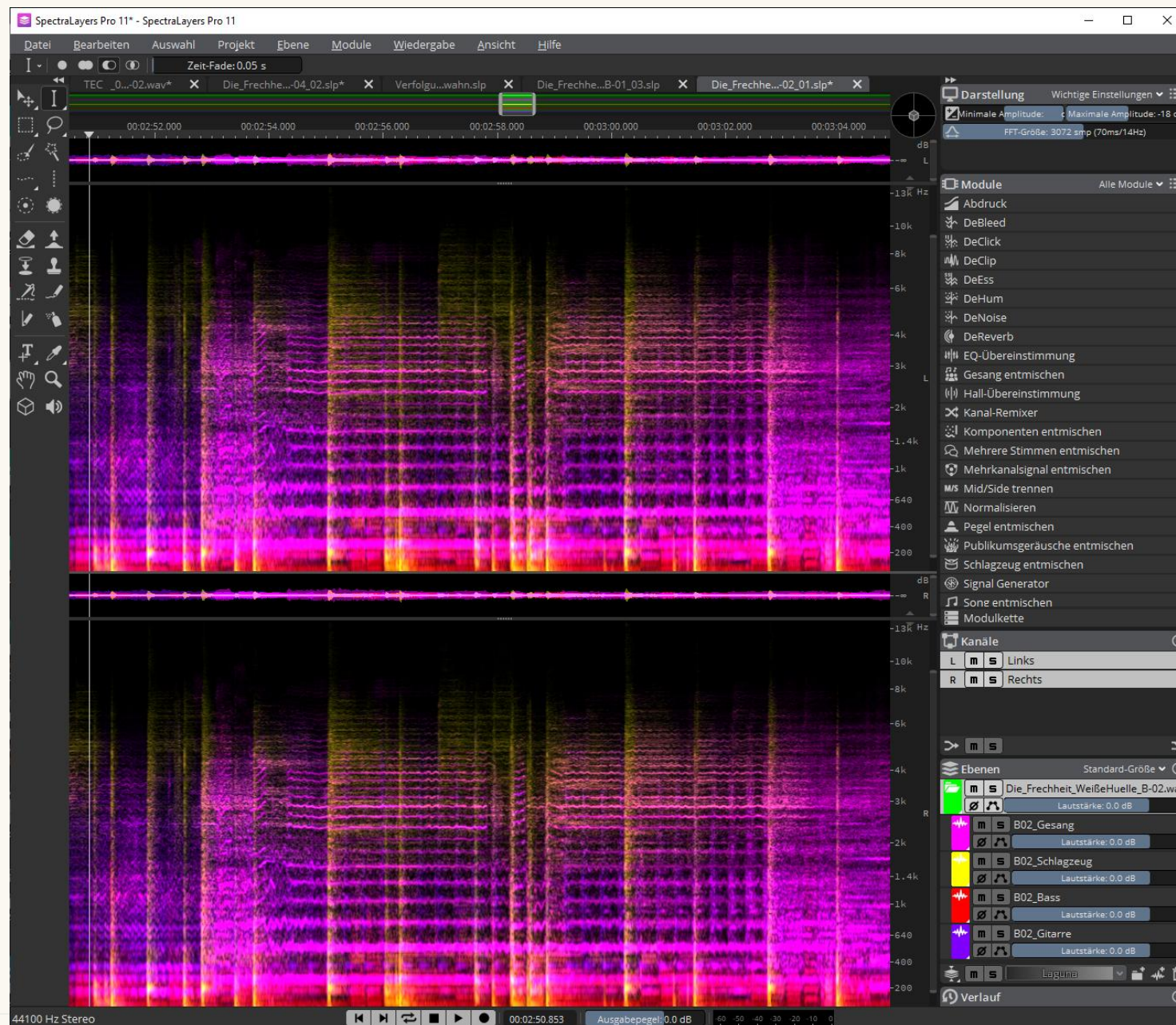


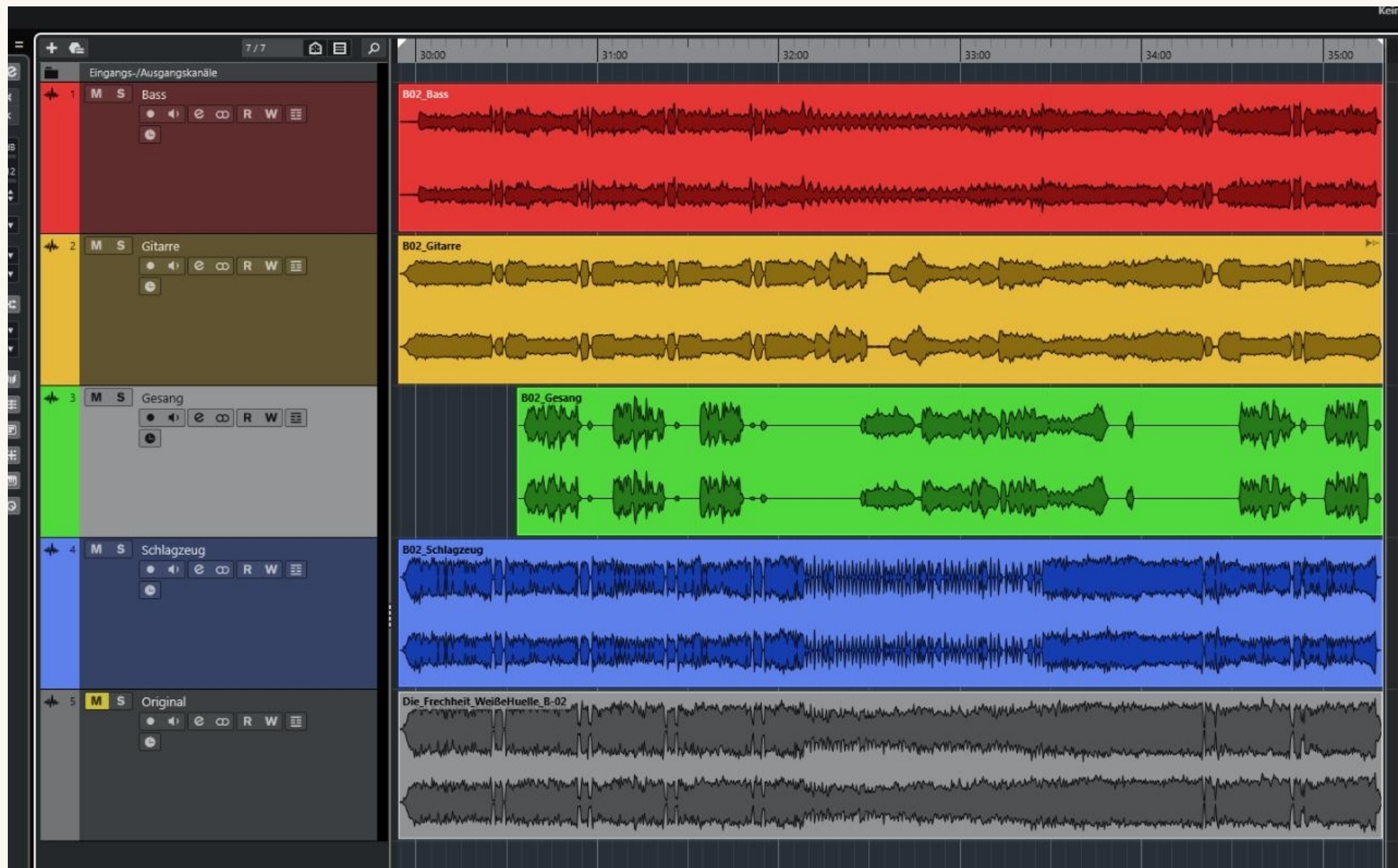






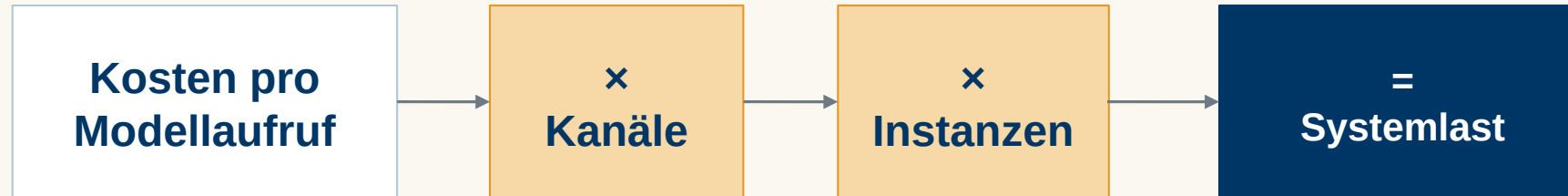






Rechenleistung: Skalierung ist die Falle

“Realtime auf einem Kanal“ ist noch keine Broadcast-Architektur.



Hardware

CPU/GPU/Accelerator

Energie

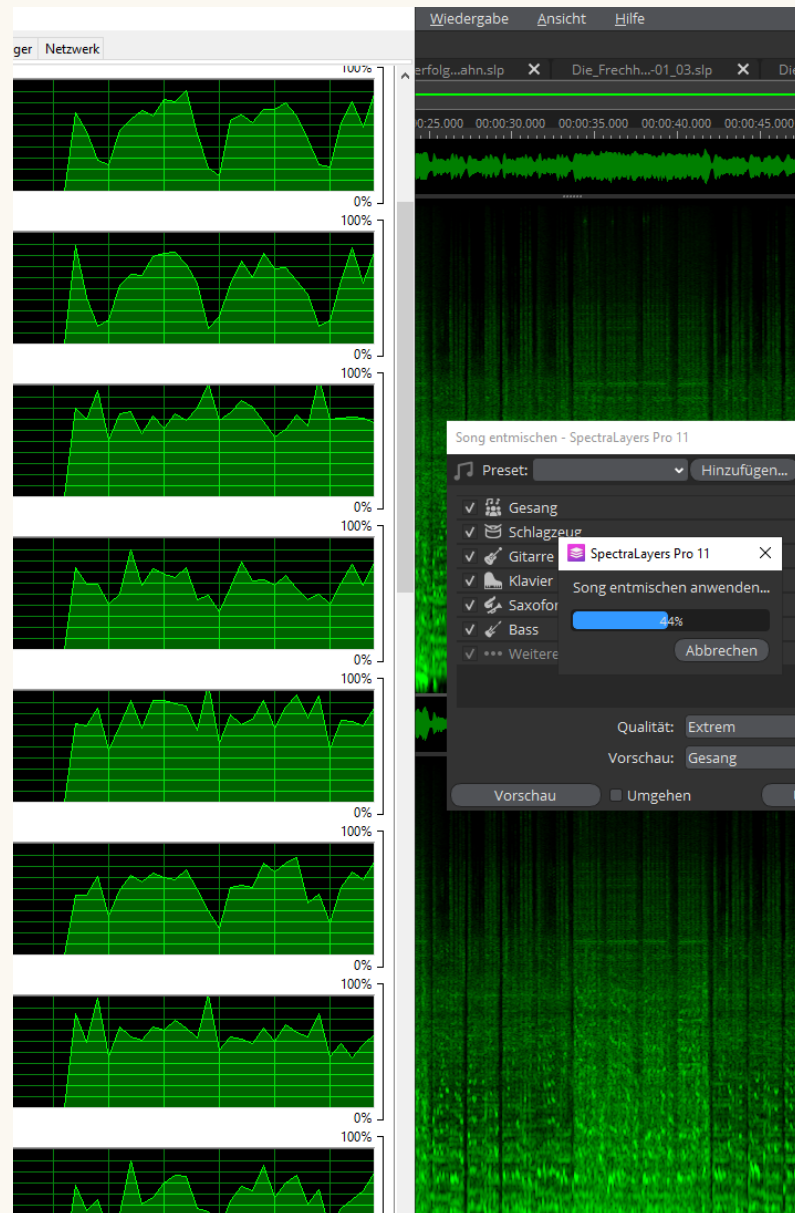
Leistung und Wärme

Dichte

Wie viele Prozesse

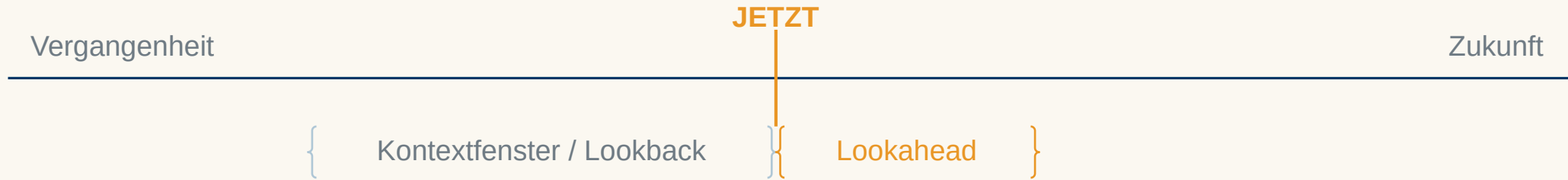
Kosten

pro System / pro Rack /
pro SaaS



Realtime und Low Latency

Ein System kann mitlaufen - und trotzdem zu spät sein



ALGORITHMUS

Fenster, Frames, Lookahead

COMPUTE

Zeit für die eigentliche Berechnung

SYSTEM

Buffer I/O, Netzwerk, Wandlungen

Im Live-Betrieb ist “Zukunft” eine knappe Ressource

DSP oder ML? Falsche Frage!

Nicht „moderner“, sondern: welche Technologie kann was?

KLASSISCHE SIGNALVERARBEITUNG

stark bei:

- exakt definierten Aufgaben
- geringer Latenz
- geringer Last
- reproduzierbarem Verhalten
- hoher Kanaldichte

MACHINE LEARNING

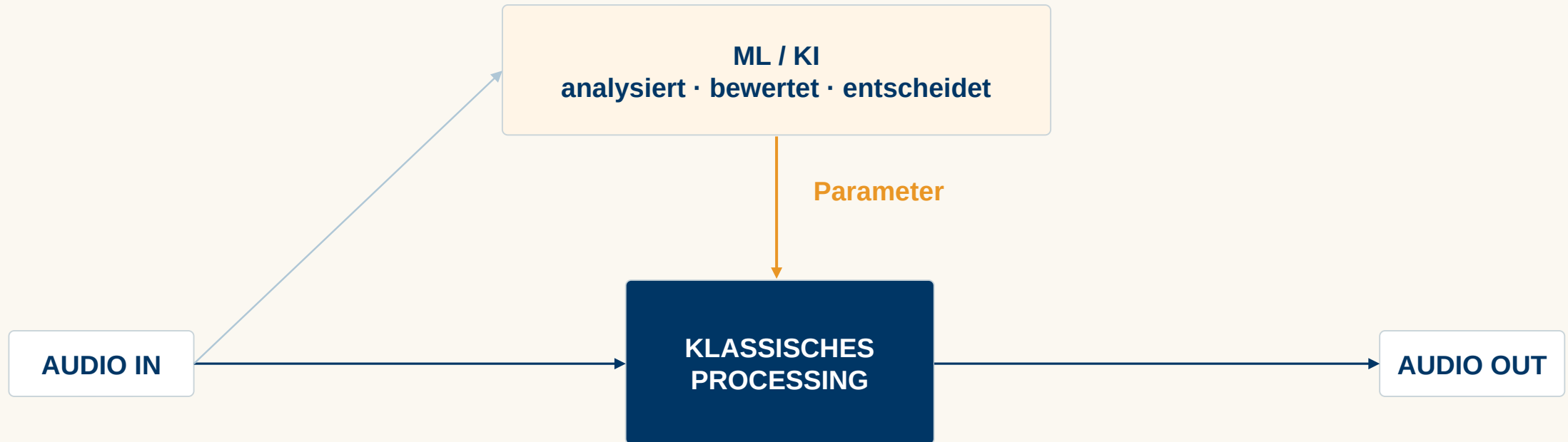
stark bei:

- Mustererkennung
- komplexen Situationen
- schwer formulierbaren Regeln
- Kontext
- Interpretation

Die richtige Architektur setzt also auf Arbeitsteilung, nicht Glaubenssätze

Die hybride Architektur

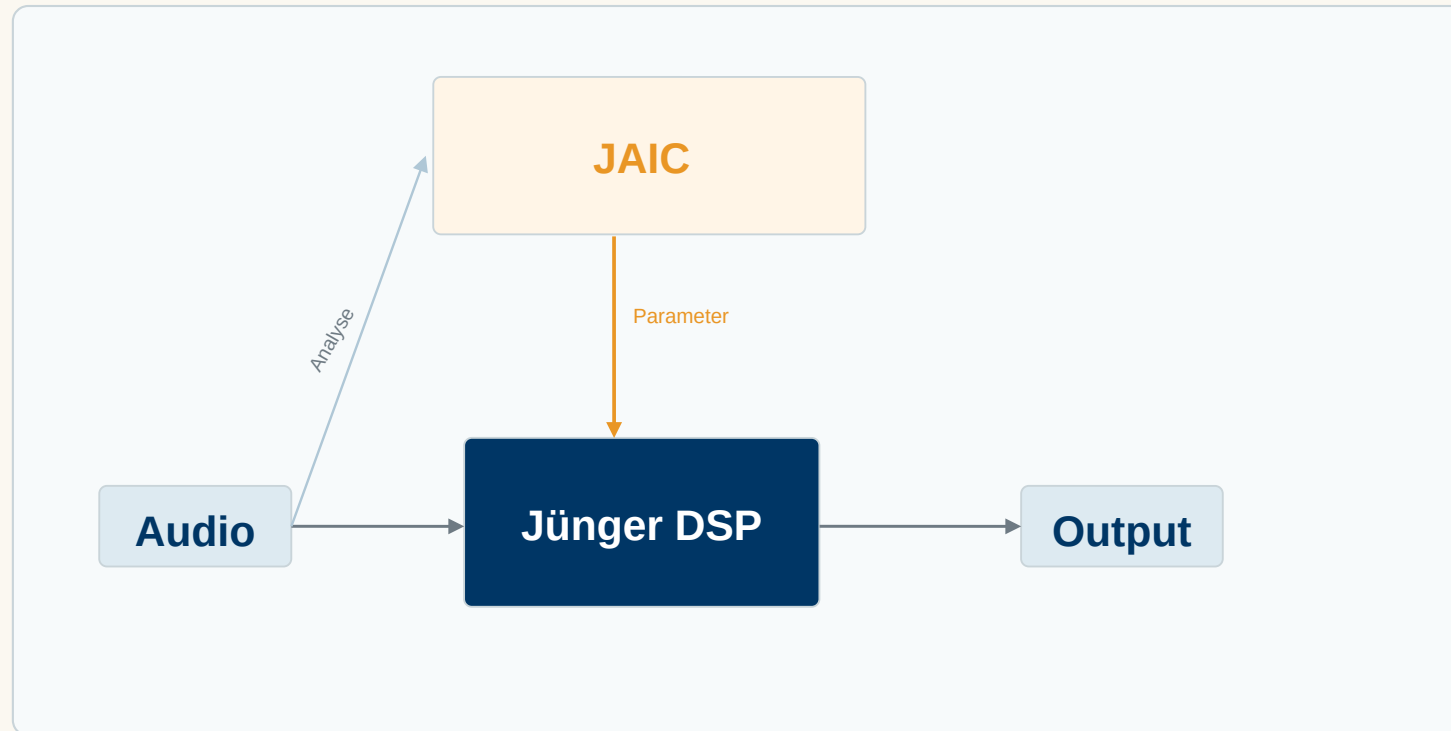
KI kann wertvoll sein, ohne selbst Samples zu verändern



Die KI sitzt eine Ebene über dem Audiopfad und überwacht

JAIC - Jünger Audio Intelligent Companion

KI unterstützt und nimmt Aufgaben ab



1 Audio kontinuierlich analysieren

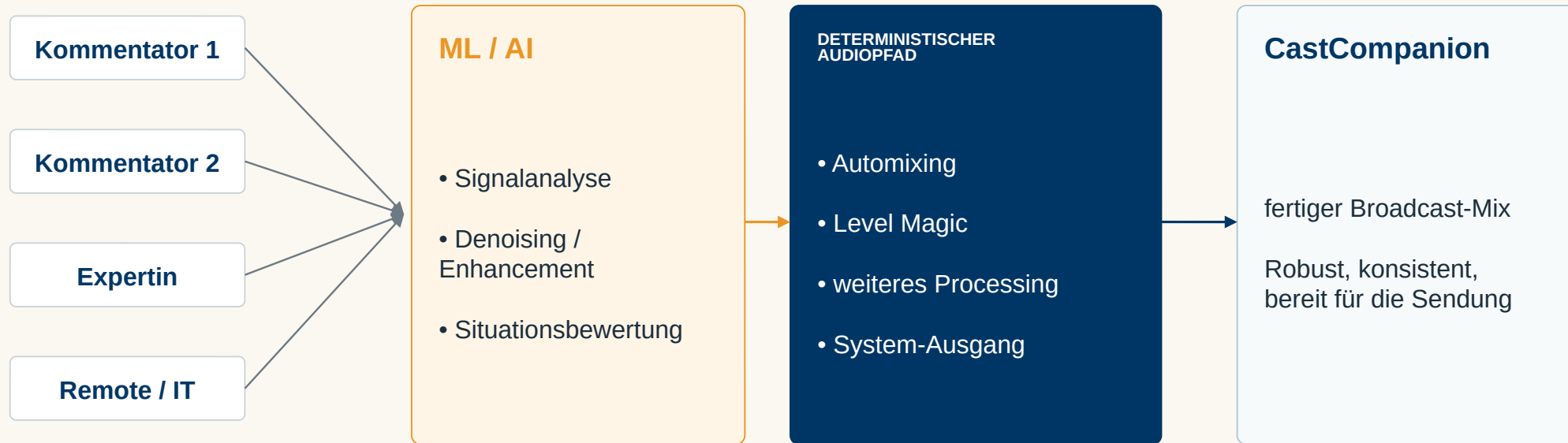
2 veränderte Situation erkennen

3 Algorithmus-Parameter anpassen

Der eigentliche Audiopfad verbleibt im kontrollierbaren Echtzeit-Processing

CastCompanion

Autonomer Voice-Mix - unterschiedliche Technologien für unterschiedliche Aufgaben

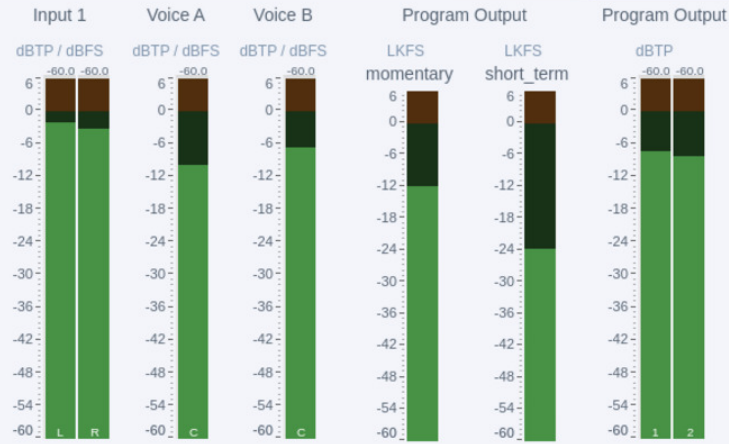


CastCompanion ist kein reines KI-Feature, sondern ein Workflow aus Analyse, Automation und bewährtem Audio-Processing.

CastCompanion - Stereo



Preset



Voice A

Noise Suppression

OFF

ON

AUTO

Pan

C

Host Priority

PRIORITIZE

Voice B

Noise Suppression

OFF

ON

AUTO

Pan

C

Host Priority

PRIORITIZE

Automix Control

Voice Prominence

BLENDED

MEDIUM

PROMINENT

ALLOW DUCKING

Loudness Leveling Mode

PROGRAM

DIALOG

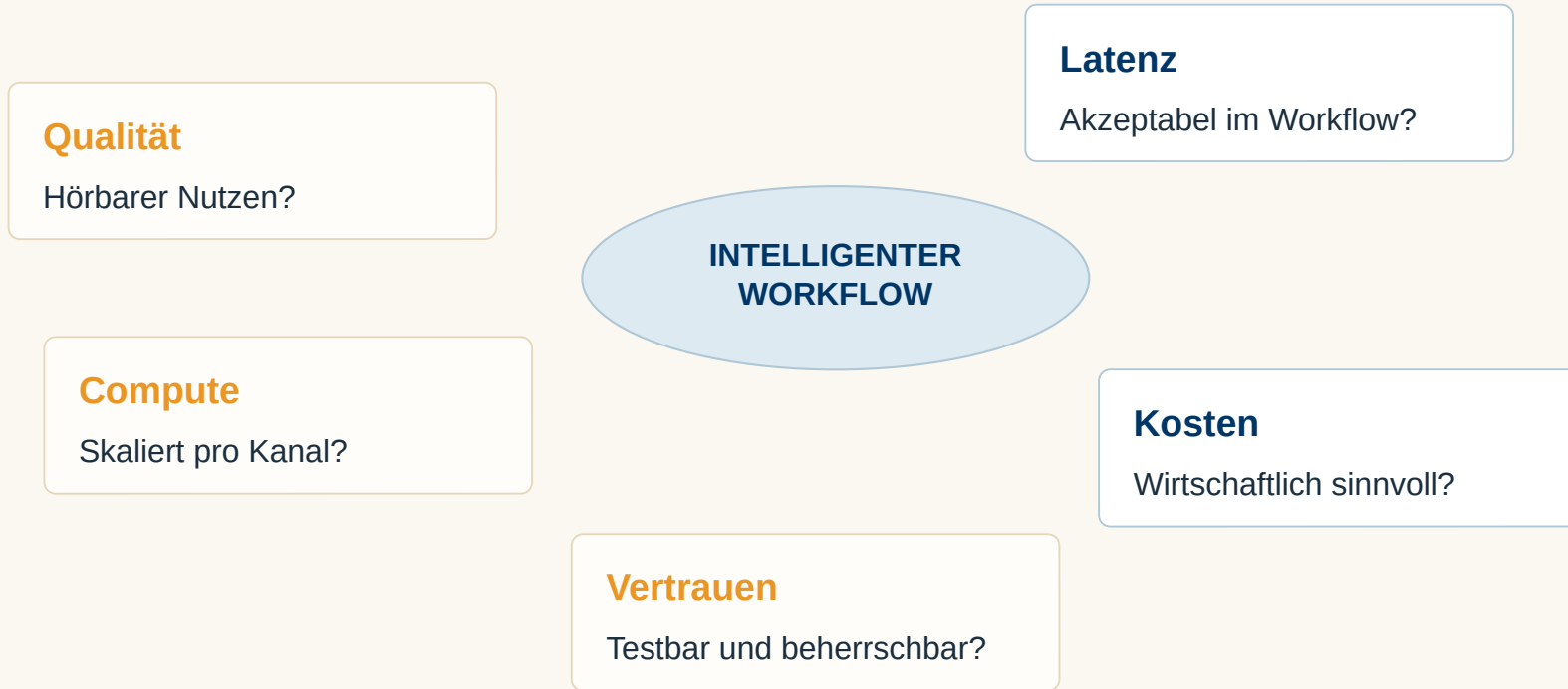
Loudness Target

-24

LKFS

Die richtige Technik am richtigen Ort

KI muss ihren Platz im Signalweg “verdienen”.



Vielleicht besteht die intelligenteste Anwendung von KI im Audioworkflow manchmal darin, sie gar nicht ans Audio zu lassen, sondern “nur” zuhören zu lassen.

Jünger Audio flexAI

