

Projekt „TranSiT KI“

Transformation von Sprache in Text
mittels künstlicher Intelligenz

**Der Einsatz von OpenAI Whisper bei der
Protokollierung von Ausschuss- und
Gremiensitzungen des Deutschen Bundestages**

Stenografischer Dienst

Arbeitsbereich „Ausschüsse und Gremien“

MRn Dr. Ursula Ortmann-Droste





Inhalt

1. Ausgangslage
2. Lösungsansatz: Sprecherunabhängige Spracherkennung und Strategieinitiative des Direktors beim Deutschen Bundestag
3. OpenAI „Whisper“
4. Projekt „TranSiT KI“
5. Umsetzung
6. Ergebnisse
7. GUIs (Faster Whisper): aTrain/noScribe
8. Ausblick
9. Workflow
10. Ein Blick in den Maschinenraum

Quelle der Icons: <https://www.flaticon.com/>



1. Ausgangslage I: Aufgaben des Stenografischen Dienstes

1. Protokollierung der Plenarsitzungen des Deutschen Bundestages

2. Protokollierung von Sitzungen von Ausschüssen und Gremien

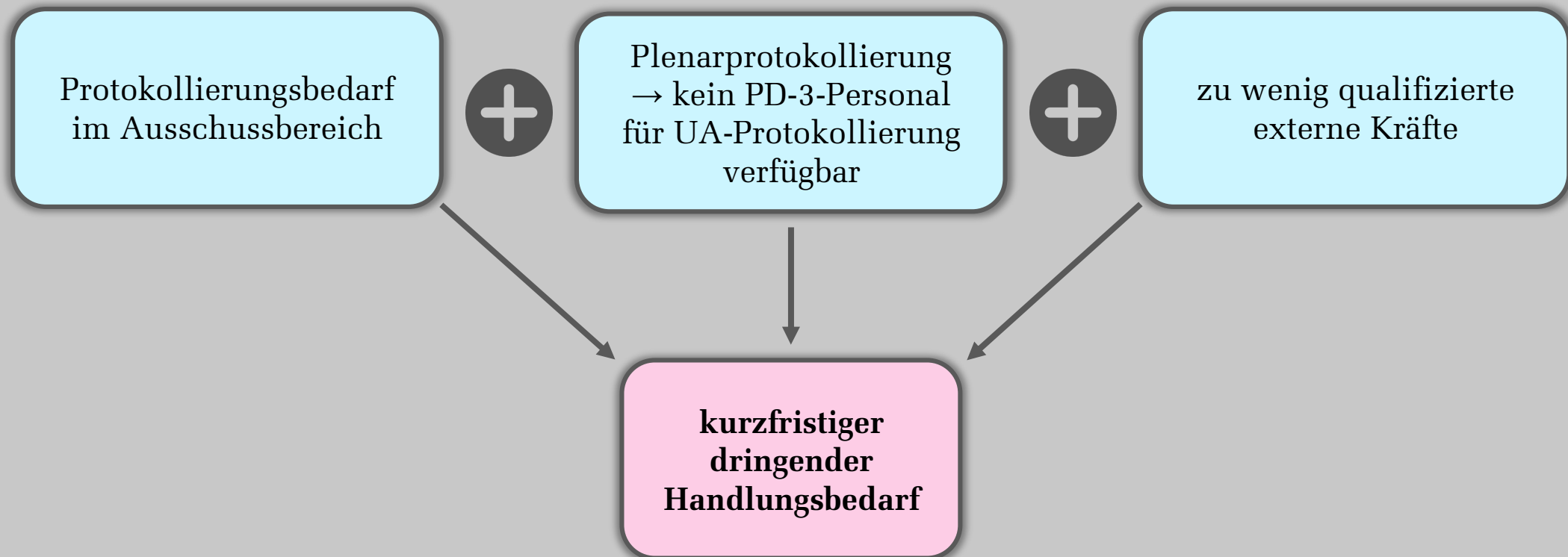
- Untersuchungsausschüsse
- Anhörungen von Fachausschüssen
- Deutsch-Französische Parlamentarische Versammlung
- Bundeswahlausschuss etc.



Sitzungen von Ausschüssen parallel zum Plenum



1. Ausgangslage II: Eigenes Personal und externe Kräfte





2. Lösungsansatz

Sprecherunabhängige Spracherkennung (Automatic Speech Recognition, ASR)

- Kein individuelles Training der Software erforderlich
(beliebige Sprecher/-innen)
- Überträgt gesprochene Sprache in einen geschriebenen Text
(Speech-to-Text-Transformation)
- Bahnbrechende Fortschritte in den letzten Jahren
- Bsp.: OpenAI Whisper

Strategieinitiative

des Direktors
beim Deutschen Bundestag
zur Modernisierung der
Bundestagsverwaltung



Dr. Michael Schäfer

3. OpenAI „Whisper“



Open Source



auf **680.000** bzw. **5 Mio.**
Std. Material trainiert



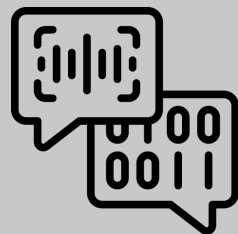
Sprachidentifizierung,
Übersetzung,
Transkription



robust gegenüber
Hintergrundgeräuschen,
Fachsprachen,
Akzenten/Dialekten



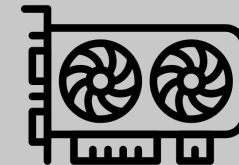
100 Sprachen



verschiedene
Sprachmodelle



online und offline verwendbar
(PD 3: Stand-alone-Notebooks
→ Datenschutz und -sicherheit)



ressourcenintensiv
→ hochleistungsfähige GPU

4. Projekt „TranSiT KI“ Details I

Ziel und Inhalt:



Insourcing outgesourcter Leistungen unter KI-Nutzung

→ Sicherstellung der Protokollierung des 1. UA

→ kurzfristige Erstellung hochwertiger Protokolle sonstiger Ausschüsse und Gremien
(auch Fremdsprachen)



Nebeneffekte

- Einsparung von Haushaltsmitteln
- Wegfall von Rekrutierungen und (Nach-)Schulungen externer Kräfte

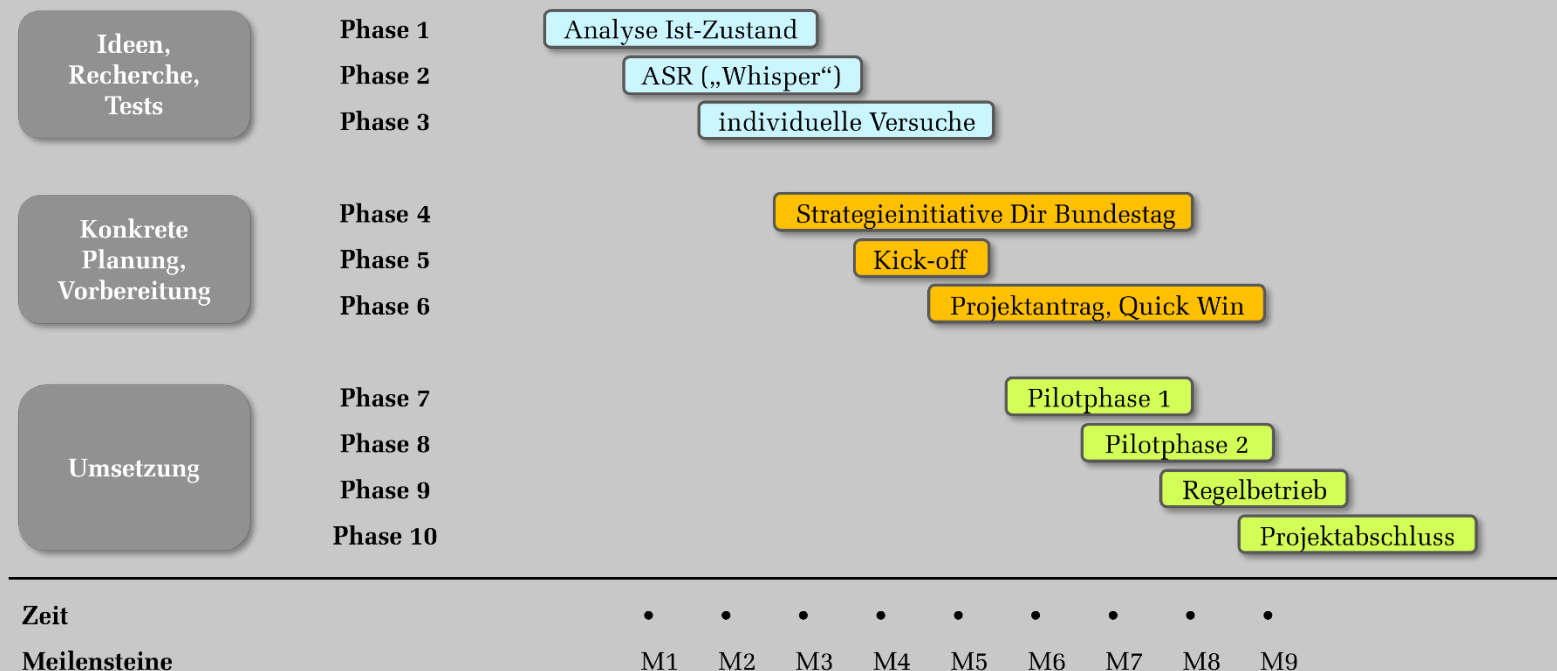


**Projektleitung, Projektteam und Projektkreis, Software-installation,
Programmierung und alleinige Verantwortung:**

Kolleginnen und Kollegen von PD 3 → Vermeidung von Reibungsverlusten

4. Projekt „TranSiT KI“ Details II

Projektphasen und Meilensteine:



5. Umsetzung I

QUICK WIN



P
I
L
O
T
P
H
A
S
E

1

- Tests
- Evaluierung und Optimierung
- Kommunikation referatsintern/mit Vorgesetzten/
anderen Org.-Einheiten

Mai-Sept. 2023



Beschaffung der „Whisper“-Notebooks

2

- Ausweitung der Tests
- Evaluierung, Optimierung und Kommunikation
- Präsentationen

Okt.-Dez. 2023

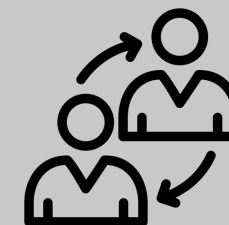


5. Umsetzung II

R
E
G
E
L
B
E
T
R
I
E
B

- Bis Januar 2025 KI-unterstützte Protokollierung
 - von mind. 50 Prozent der Sitzungen des 1. und 2. UA (über 250 Std. in dt. und engl. Sprache)
 - in weiteren Ausschuss-/Gremiensitzungen (in dt., engl., franz. Sprache)
- Fortgesetzte Evaluierung und Optimierung
- Unterstützung anderer Parlamentsverwaltungen (national und international)
- Austausch mit dem BSI
- Teilnahme an internationalen Konferenzen, Input
- Wissenschaftlicher Aufsatz zum Einsatz von „Whisper“
- Interview mit dpa
- Koordinierung bei Beantwortung von Presseanfragen und Zuarbeit

seit Jan. 2024



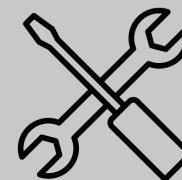
6. Ergebnisse

„Whisper“



Qualität

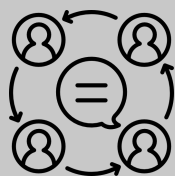
- Deutsch: sehr gut (inkl. Interpunktion)
- schlechte Tonqualität, undeutliche Aussprache, starker Dialekt: gut bis akzeptabel
- Englisch: hervorragend



Verbesserungs-
fähig

- Sprechererkennung (Speaker Diarization)
- Sprachwechsel (mehrsprachige Sitzungen)

Referatsinterne Kommunikation



- Einbeziehung aller Dienstangehörigen
- positives Feedback auf allen Ebenen

Resonanz



- PD 3 ist gesuchter Gesprächspartner
- Expertise wird (inter)national nachgefragt
- erhebliche Einsparungen von Haushaltsmitteln



7. GUIs (Faster Whisper): aTrain/noScribe



aTrain

- Open Source
- Business Analytics and Data Science-Center, Universität Graz
- Faster Whisper + Pyannote.audio
- Installation auch über Microsoft App-Store
- vollständig offline
- Sprechererkennung
nur sehr bedingt zuverlässig

noScribe

AI-powered Audio Transcription

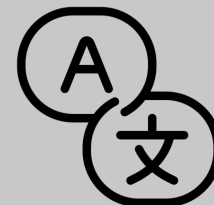
- Open Source
- Universität Luzern (Kai Dröge) und Institut für Sozialforschung Frankfurt/Main
- Faster Whisper + Pyannote.audio
- Installationsdatei über SWITCHdrive
- vollständig offline
- Sprechererkennung
nur sehr bedingt zuverlässig



8. Ausblick

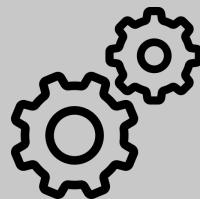


Verbesserungen bei der
Sprechererkennung



Verbesserungen bei Sprachwechseln
(bei mehrsprachigen Protokollen)

Weiterhin operatives
Prozessmanagement



„Whisper“-Einsatz
bei Protokollierung eingestufter Sitzungen



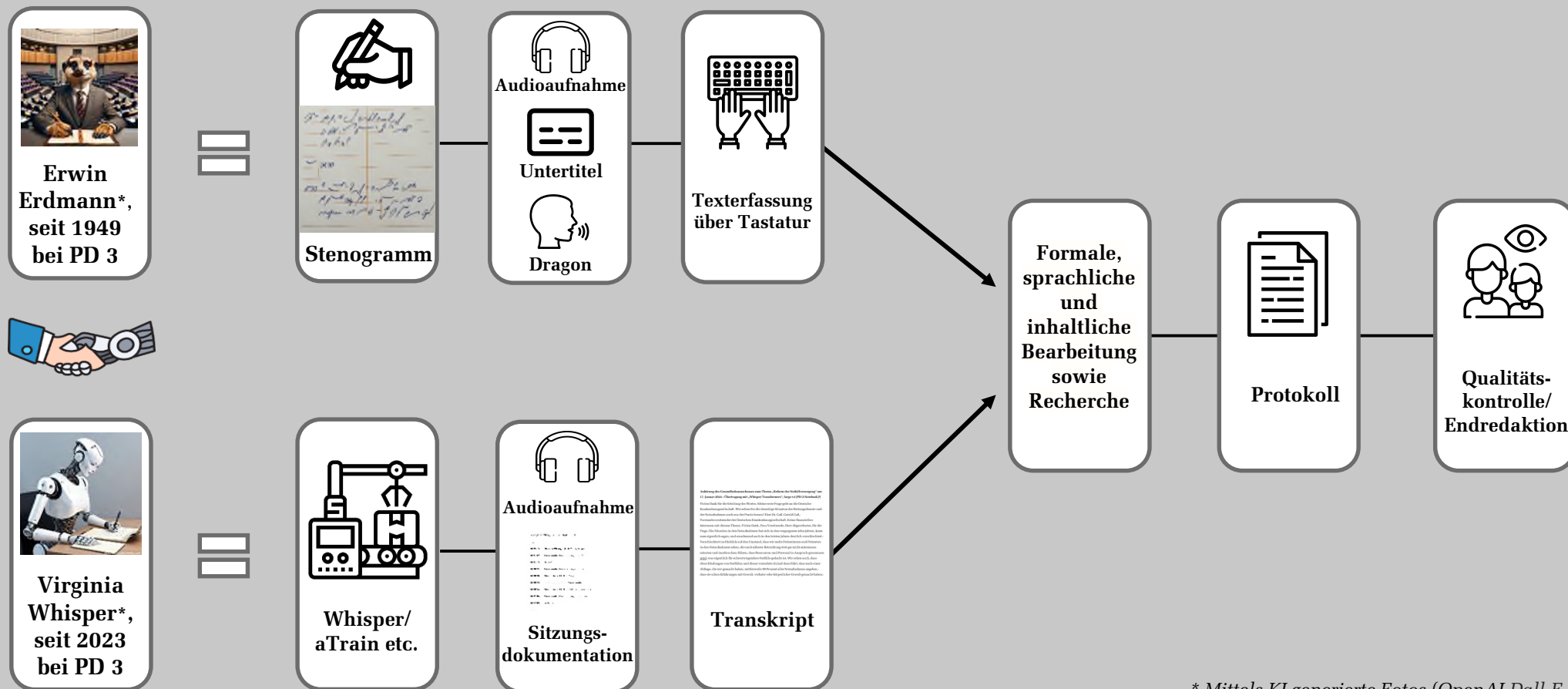
**Innovations-
projekt**

Training von „Whisper“
auf UA-spezifische sprachliche
Besonderheiten



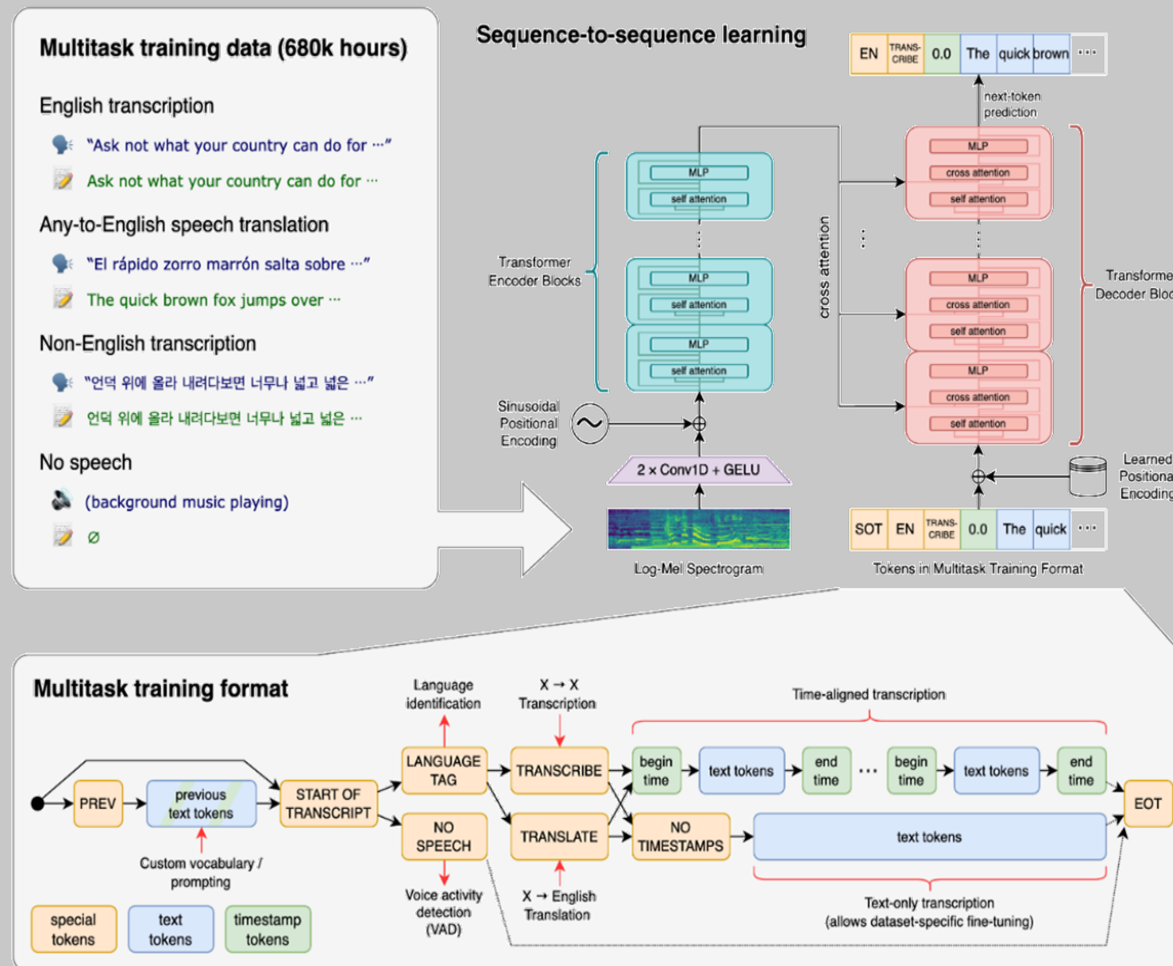
Kommunikations-/Beratungsleistungen
(intern und extern)

9. Workflow



* Mittels KI generierte Fotos (OpenAI Dall-E 3 und Stable Diffusion)

10. Ein Blick in den Maschinenraum I

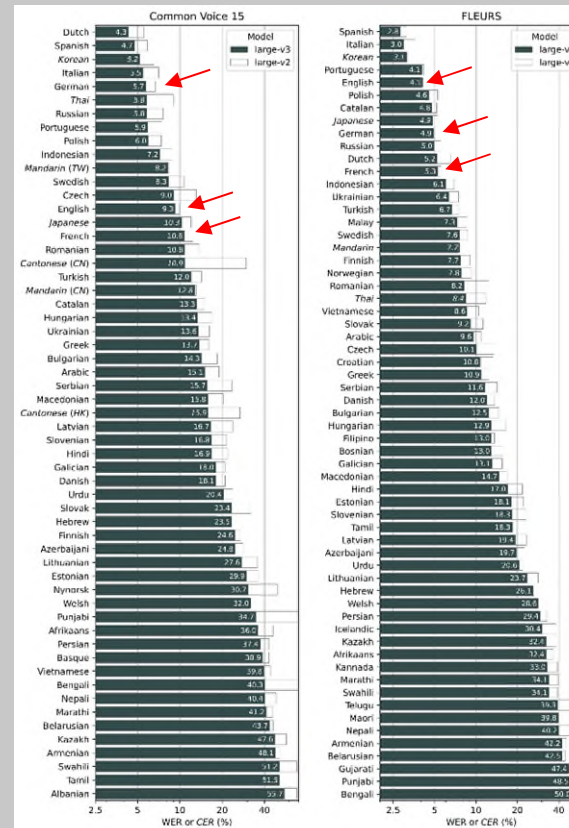


Transformer Sequence-to-sequence-Modell
(<https://github.com/openai/whisper>)



10. Ein Blick in den Maschinenraum II

Fehlerquoten (Word Error Rates, WERs) und Character Error Rates, CERs) bei den Sprachmodellen large-v3 und large-v2 in Abhängigkeit von Testdatensätzen (Common Voice 15 und Fleurs)



(<https://github.com/openai/whisper> - Bearbeitung: Stenografischer Dienst)

10. Ein Blick in den Maschinenraum III



Stand-alone-Notebook PD 3 mit Jupyter Notebook

Spezifikationen: GPU: NVIDIA® GeForce RTX 4090,
16 GB VRAM, CPU: Intel® Core i9-13900HX, 64 GB RAM,
SSD-HD mit 1 TB Speicher, Windows 11 Pro

Foto: Stenografischer Dienst



Erwin Erdmann*,
seit 1949 bei PD 3



Virginia Whisper*,
seit 2023 bei PD 3

**Vielen Dank für Ihr Interesse und
Ihre Aufmerksamkeit!**



** Mittels KI generierte Fotos (OpenAI Dall-E 3
und Stable Diffusion)*