

# A Privacy-Aware Web Application for Swiss German and Multilingual Speech Transcription

Institute of Computer Science (InIT)  
17.06.2025 | Flavio D'Intino

# Team



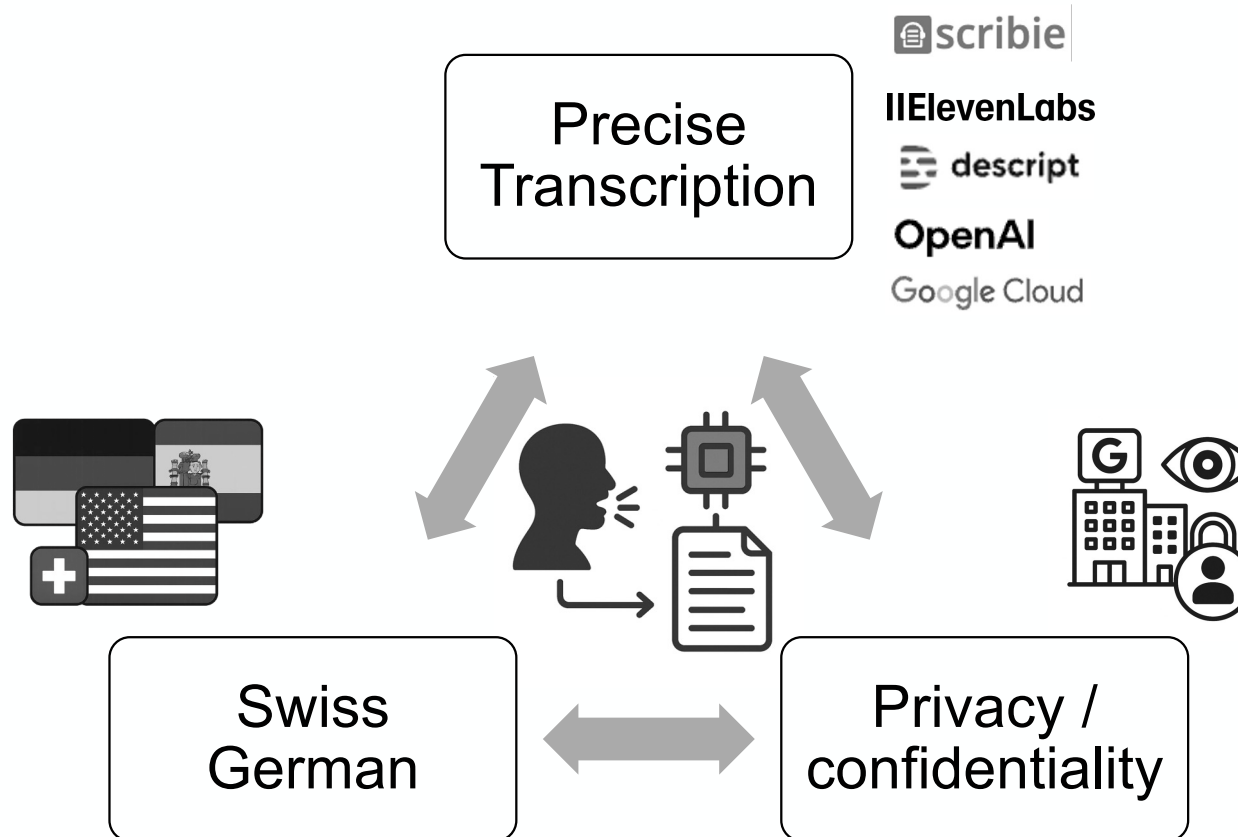
Flavio D'Intino  
Research Associate  
[dint@zhaw.ch](mailto:dint@zhaw.ch)



Hans-Peter Hutter  
Head of Research Group Human-Centered Computing  
[huhp@zhaw.ch](mailto:huhp@zhaw.ch)



# Context



# Swiss German Speech-to-Text

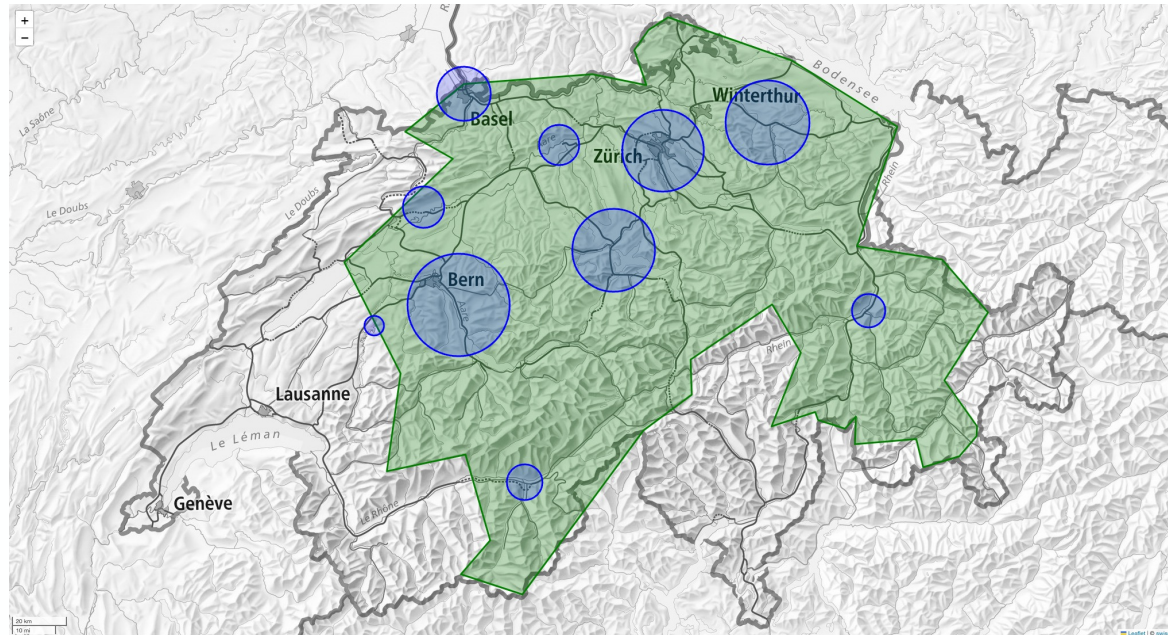
- Controlled speech: 14 - 25% WER
- Spontaneous speech: 30 - 40% WER
- Challenges
  - Low resource
  - STT & translation to Standard German
  - 100s of dialects
  - No standards
  - Mixed-lingual
    - Swiss German
    - French
    - Italian
    - English

# Model Training with Benchmark Datasets

- Controlled recordings
- Quiet environments
- Few spontaneous speeches
  - Translated from Standard German
  - Read from a manuscript
  - No disfluencies, hesitations, slip of the tongue, repetitions, etc.
- Single speaker per sample
- Short samples: single utterance

→ Collect new realistic SRB-300 Dataset

# SRB-300 Speech Corpus



# Comparison

## Radio Rottu Oberwallis

**Our** (5.1% WER, 93.6 BLEU):

Das ist ein klares Bekenntnis, dass alle vom Tourismusverein, da sind ja alle Leistungsträger drin, **G-Shield**, Hotelier, Gruppenhäuser, Apartments, Bergführer, die sich klar bekennen, im Saas braucht es die Hannigbahn. Dass sich der Verein am Neubau beteiligen könnte, ist noch im Herbst nicht überall im Tal gut angekommen. Auf kritische Stimmen angesprochen, sagt Konstantin Baumann.

**Reference System** (26.7% WER, 72.0 BLEU):

Das ist ein klares Bekenntnis, dass alle vom Tourismusverein da sind ja alle Leistungsträger **d. Skischil**, die Hotelier, **Gruppe hier ist e** Apartment, Bergführer, die sich klar bekennen, im Saas braucht es die **Handigbahn**. Dass sich der Verein am Neubau beteiligen könnte, ist noch am Herbst nicht überall im Tal gut angekommen. Auf kritische Stimmen angesprochen, sagt Konstantin Bumann.

**Ground truth:**

Das ist ein klares Bekenntnis, dass alle vom Tourismusverein, da sind ja alle Leistungsträger **drin, Skischulen**, Hotelier, **Gruppenhäuser**, Apartments, Bergführer, die sich klar bekennen, im Saas braucht es die **Hannigbahn**. Dass sich der Verein am Neubau beteiligen könnte, ist noch im Herbst nicht überall im Tal gut angekommen. Auf kritische Stimmen angesprochen, sagt Konstantin Bumann.

10 min sub-set	WER (%)	BLEU
Competitor system	20.0	70.5
Our	<b>16.5</b>	<b>71.2</b>



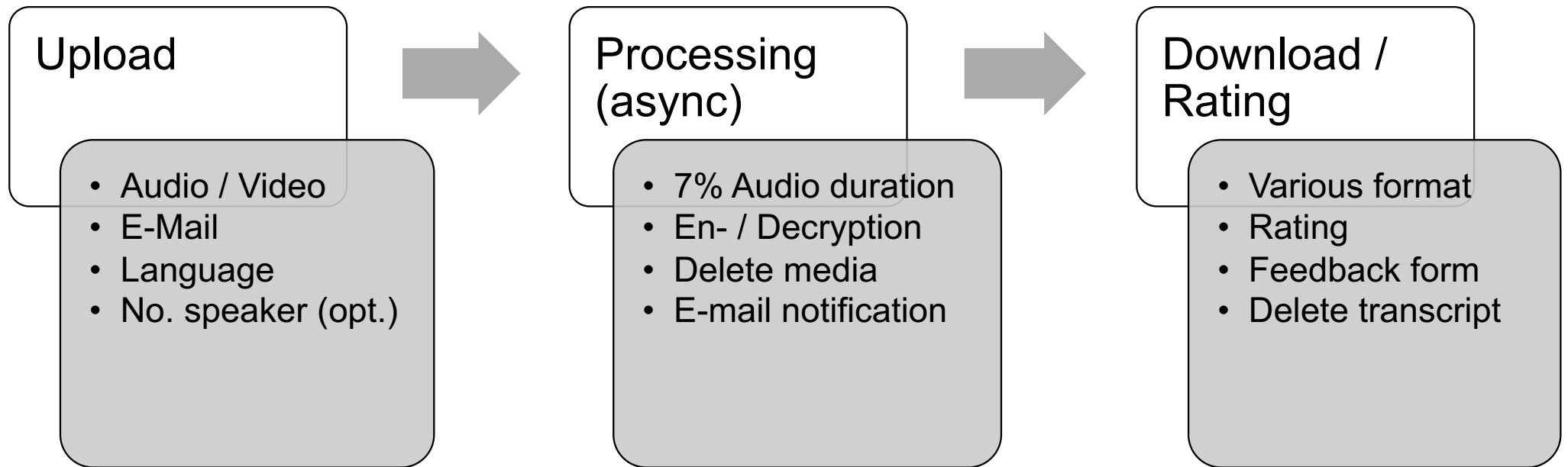
Overall	WER (%)	BLEU
Ref. zero-shot	44.2	37.5
Reference system	23.6	62.8
Zero-shot	20.8	70.9
Our	<b>16.1</b>	<b>75.9</b>

# Web Service

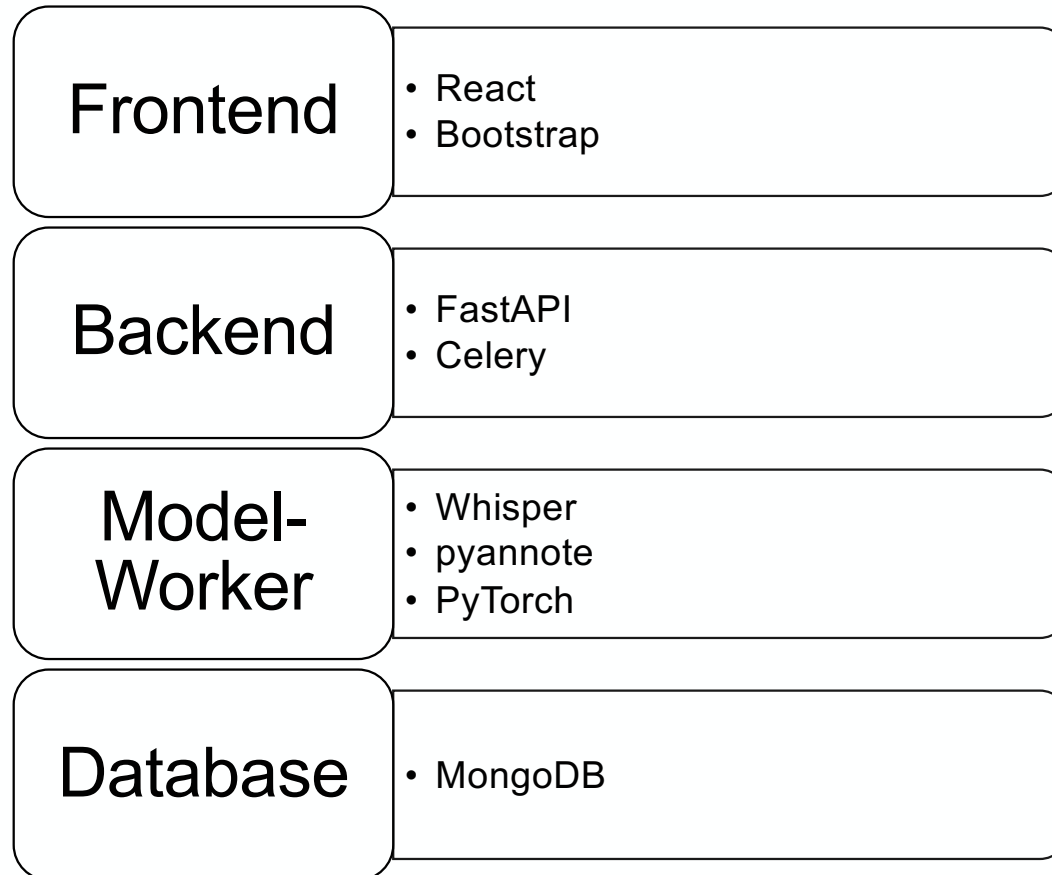
The screenshot shows the main upload page of the 'Transkriptions-Service für Schweizerdeutsch'. The browser address bar shows 'transcription.init-lab.ch/#/'. The page header includes the 'zhaw' logo, 'Transkriptions-Service', navigation links for 'Home', 'FAQ', and 'Login', and the text 'Institut für Informatik'. The main content area features a heading 'Transkriptions-Service für Schweizerdeutsch' followed by a description: 'Dieser Service erkennt Schweizer Dialekte und transkribiert sie in Standardhochdeutsch. Weitere Sprachen werden ebenfalls unterstützt.' Below this, it states 'Ungefähre Wartezeit: 1 min'. There is a form for 'E-Mail Adresse' with the placeholder 'e-mail (zhaw.ch)'. The 'Audio/Video Datei' section has a 'Choose file' button and a 'No file chosen' status. A note lists supported file types: '.aac, .flac, .m4a, .mp3, .mp4, .wav, .wave, .wma' with a maximum file size of 700 MB. The 'Sprache' dropdown is set to 'Schweizerdeutsch'. There is an 'Anzahl Sprecher (optional)' input field and an 'Upload' button. A 'Datenschutzerklärung' section explains that audio and transcription files are encrypted and deleted after 24 hours, with a link to the full disclaimer. The footer contains links for 'Kontakt', 'Über uns', 'Disclaimer', and '© 2025 InIT'.

The screenshot shows the download and evaluation page. The browser address bar shows 'https://transcription.init-lab.ch/#/download-and-rate/42391f4e...'. The page header is identical to the previous screenshot. The main content area features a heading 'Download und Bewertung von d06cb051b5054e4599a58ae5872b883c.mp3'. Below the heading, it says 'Viele Dank für die Nutzung unseres Service. Sie können die Transkription hier herunterladen.' There are five buttons for download options: '.txt', '.tsv', 'MAXQDA', 'MAXQDA (kompakt)', and '.srt'. A note says 'Fahren Sie über die Buttons für weitere Infos.' Below this, it says 'Wir würden uns über eine Rückmeldung freuen. Nehmen Sie sich ein paar Augenblicke Zeit.' There are two rows of five smiley face icons for rating the transcription quality and the application quality. The first row (transcription quality) has three sad faces and two happy faces. The second row (application quality) has two sad faces and three happy faces. There is a 'Kommentar:' section with a text area containing the placeholder 'Verbesserungen, Probleme, ...' and a 'Bewertung abgeben' button. The footer is identical to the previous screenshot.

# Web Service



# Tech Stack



# Features

- Import to MAXQDA
- Speaker diarization
- Multi-lingual
- Download in various formats
- Processing & data storage on our campus

## **Future**

- Fine-tuning on domain-specific vocabulary
- Download wizard

# Applications at ZHAW

## **Marketing & Communications**

- Video subtitles
- Interview transcriptions

## **Schools of Applied Psychology, Health Sciences, Linguistics, and Management & Law**

- Confidential interview transcription
- Group discussion transcription

## **Our industrial partner**

- Automatic minutes of meetings, incl. timestamped full transcription

# Outlook

- 25th June: «Advancing STT for Low-Resource Real-World Speech» at HCII 2025
- Scale
  - Onboard further institutions?
  - Commercialization?





Questions or inquiries?

Feel free to contact us at  
[dint@zhaw.ch](mailto:dint@zhaw.ch)

We thank Innosuisse for their valuable support in making this project happen.

Zurich University of  
Applied Sciences

**ZHAW School of Engineering**  
Institute of Computer Science

Flavio D'Intino  
Research Associate  
Steinberggasse 13  
CH-8401 Winterthur  
dint@zhaw.ch  
www.zhaw.ch/init