

# Evaluating and Enhancing Speech To Text Models for Industry Specific Jargon

Leonhard Stengel

17.04.2025, Kick-off Presentation

Chair of Software Engineering for Business Information Systems (sebis)  
Department of Computer Science  
School of Computation, Information and Technology (CIT)  
Technical University of Munich (TUM)  
[www.matthes.in.tum.de](http://www.matthes.in.tum.de)

Motivation

Research Questions

Approach

Timeline

## SAP Joule - KI Assistant integrated into SAP Software

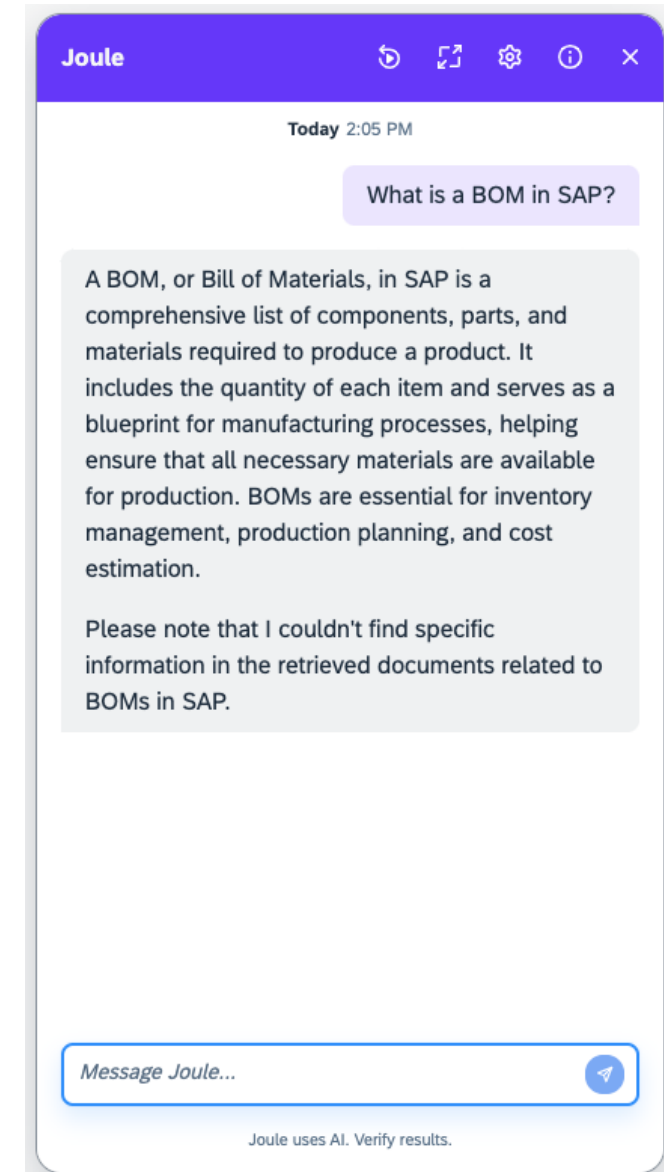
Joule should also support **speech input** in the future

- Faster and more efficient way of input
- Hands free interaction

In a first evaluation SAP found that state of the art speech to text models have problems in transcribing words of industry specific jargon

These results correspond to current research

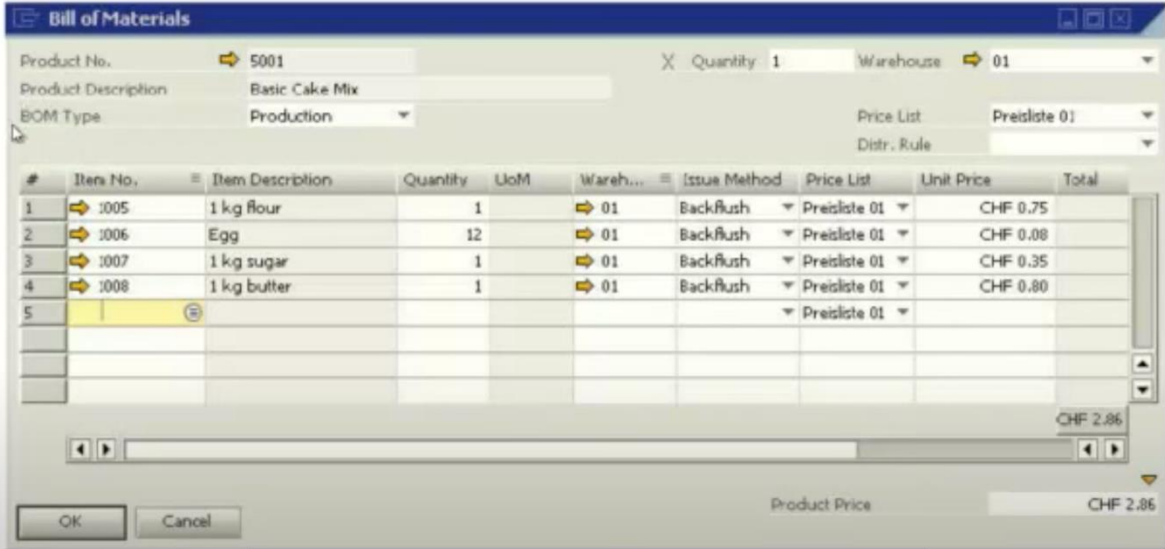
- Nguyen et al. (2024)
- Shamsian et al. (2024)



# Industry Specific Jargon

Jargon refers to a set of tokens that is uncommon with out-of-domain people

<https://www.youtube.com/watch?v=mK-qW2cz0pU>



and see the bill of materials referred to as a bomb at transalta you will

Basic SAP Terms Part 2 - Master Data

bomb  $\neq$  BOM

<https://www.youtube.com/watch?v=aW2LvQUcwqc>



normal Lotus o deltoid type placed in panand dermic semi bolid Slots

Technical Jargon Overload

YouTube automatic captioning fails quite often

Motivation

Research Questions

Approach

Timeline

RQ 1

How to benchmark out-of-domain performance of speech to text models (STT) based on industry specific jargon?

RQ 2

Which metrics enable the evaluation of out-of-domain vocabulary in STT models?

RQ 3

How can the out of domain performance of STT models be improved on industry specific jargon?

RQ 1

How can the out-of-domain performance of STT models be systematically benchmarked on industry specific jargon?

RQ 2

Which evaluation metrics best capture the recognition performance of STT models on domain specific vocabulary?

RQ 3

What techniques can improve the recognition of industry specific jargon in out-of-domain STT scenarios?

# Outline



Motivation

Research Questions

Approach

Timeline

## RQ 1

How can the out-of-domain performance of STT models be systematically benchmarked on industry specific jargon?

### Goals:

- Design an evaluation pipeline
- Find and create suitable datasets

### First findings:

- SAP training videos with transcript
- Existing jargon-datasets: Medical, Voxpopuli, ...

## RQ 2

Which evaluation metrics best capture the recognition performance of STT models on domain specific vocabulary?

### Goals:

- Find suitable metrics for evaluation
- Implement evaluation pipeline
- Evaluate state of the art STT models

### First findings:

- Whisper, aiOla

## RQ 3

What techniques can improve the recognition of industry specific jargon in out-of-domain STT scenarios?

### Goals:

- Find current research on out-of-domain STT techniques
- Implement techniques
- Evaluate results

### First findings:

- Jargon Injection
- Keyword-Guided Adaptation

## Key Idea:

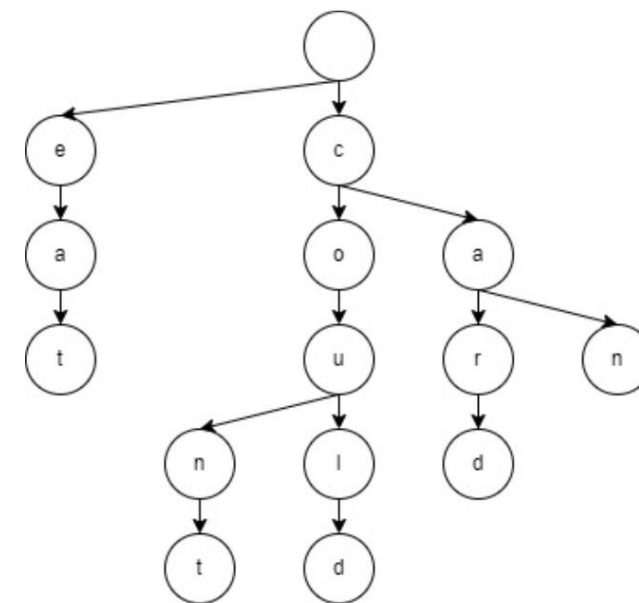
- Represent jargon as a trie tree
- Guide decoding using modified beam search to boost beams that contain jargon phrases
- Example: "insuline" (rare) vs "instruction" (common). When detected in trie tree boost beam of "insuline".

## Benefits:

- Lightweight (no fine-tuning needed)
- Works with encoder-decoder based models (e.g. whisper)

## Relevance to Thesis

- Method to improve STT output on industry specific vocabulary
- Works without labeled training data
- but the whole vocabulary needs to be in the trie tree



Trie tree: enable fast, incremental match lookup

## Key Idea:

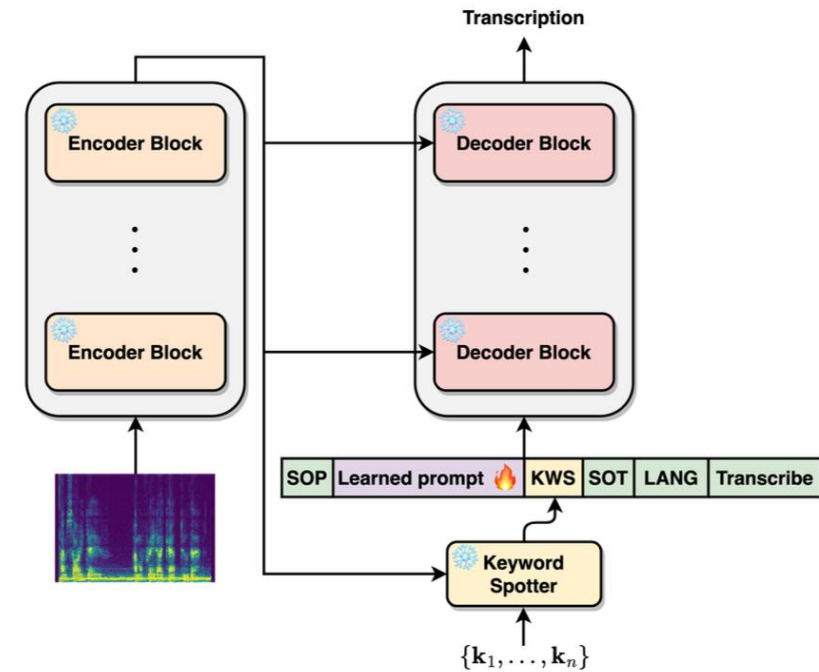
- **Keyword Spotting** (AdaKWS) detects terms from audio
- Keywords are turned into a prompt Inject
- Inject prompt into decoder

## Advantages:

- Works with small datasets
- Good generalization to new domains

## Relevance to Thesis

- Provides a trainable biasing approach
- Fits low-data scenarios



# Outline

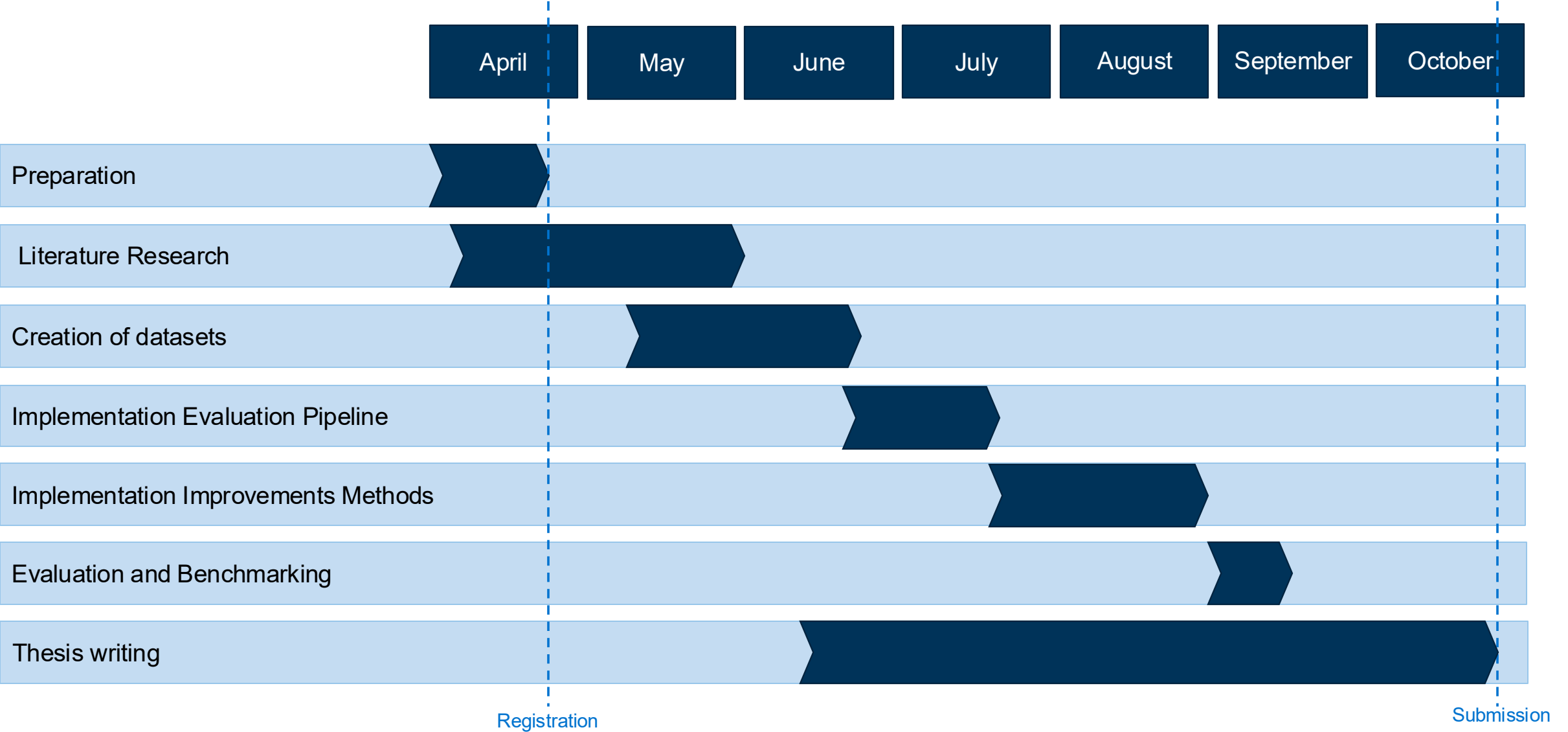


Motivation

Research Questions

Approach

Timeline





Prof. Dr.

**Florian Matthes**

Technical University of Munich (TUM)  
TUM School of CIT  
Department of Computer Science (CS)  
Chair of Software Engineering for Business  
Information Systems (sebis)

Boltzmannstraße 3  
85748 Garching bei München

+49.89.289.17132  
matthes@in.tum.de  
[www.matthes.in.tum.de](http://www.matthes.in.tum.de)

