SELF-MADE-ATC

SELF-MAINTAINING DIGITAL EVOLUTION FOR AIR TRAFFIC CONTROL SUPPORT

CURRENT STATUS OF TOOL DEVELOPMENT

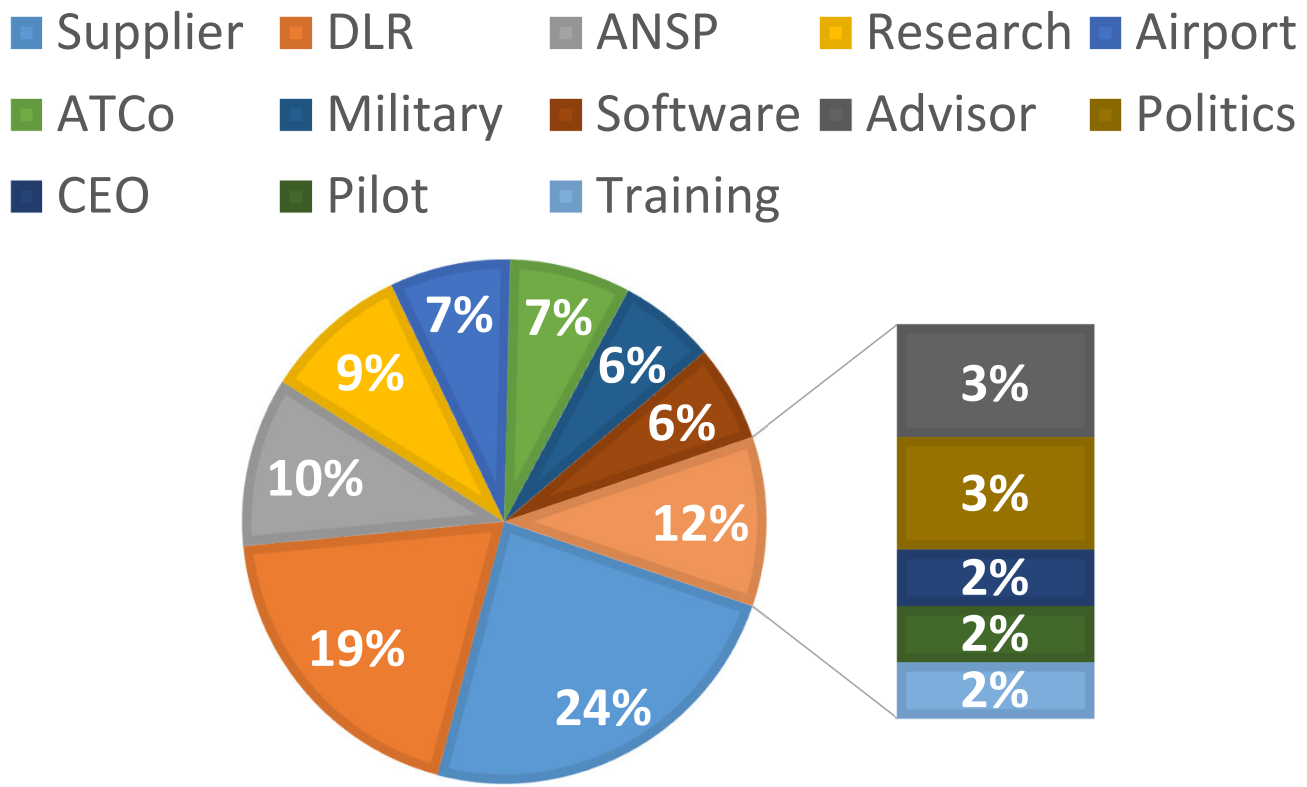
Prof. Dr. Hartmut Helmke

SELF-MADE-ATC

DLR

Welcome to Braunschweig

- Approx 65 attendants, 20 nations, 4 continents



Welcome to Braunschweig

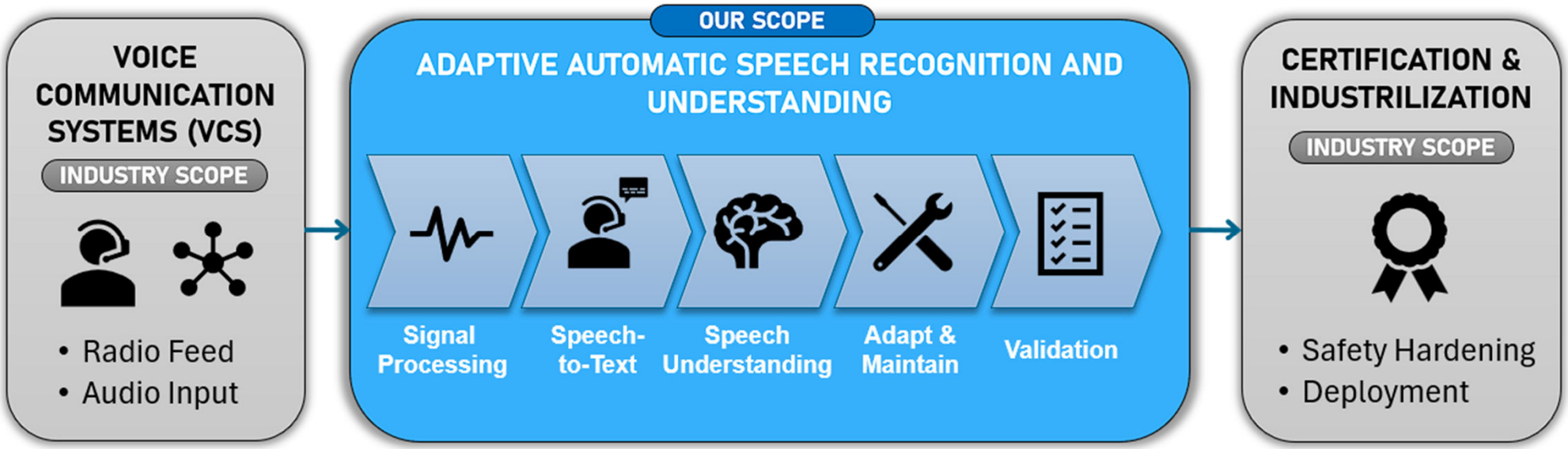
- Many arrived by plane, but without usage of Automatic Speech Recognition.

- We have Alexa, Siri, …
- Let's benefit from Automatic Speech Recognition also in ATM.
- Speech was a game changer in the development of humankind.

SELF-MADE-ATC helps to automatically adapt a Speech Recognition and Understanding application developed for one airport, e.g. Vienna, to another airport, e.g. Prague.

It is a complete tool-suite, not just a concept.

Why SELF-MADE-ATC? DLR Covers the Whole Chain



- Many components are research prototypes
- Learn, where improved prototypes are needed and what you already have.
- Let's not focus on problems, they are challenges. Let's focus on the solutions.



Overview today

1. Motivation for SELF-MADE-ATC
2. Speech Recognition and Speech Understanding
3. Building Blocks
   a. Data Collection,
   b. Model Train,
   c. Adapt,
   d. Maintain,
   e. Evaluation,
   f. Monitor
4. Conclusions
5. Overview of today and tomorrow

SELF-MAINTAINING DIGITAL EVOLUTION FOR AIR TRAFFIC CONTROL SUPPORT

3:55



1. **Motivation for SELF-MADE-ATC**
2. Speech Recognition and Speech Understanding
3. Building Blocks
   a. Data Collection,
   b. Model Train,
   c. Adapt,
   d. Maintain,
   e. Evaluation,
   f. Monitor
4. Conclusions
5. Overview of today and tomorrow

SELF-MAINTAINING DIGITAL EVOLUTION FOR AIR TRAFFIC CONTROL SUPPORT

Automatic Speech Recognition and Understanding (ASRU)

ASRU adapted for Vienna Approach 24 person months of effort

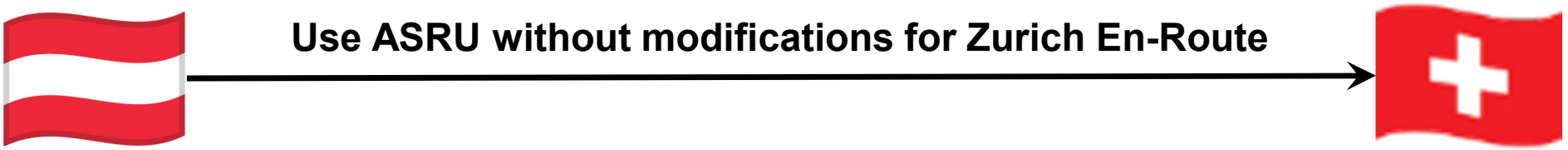


austrian four seven two tango
good morning wien radar
identified follow balad three
november transition



Performance:
Word Error Rate = 3%
Command Recognition = 92%
Callsign Recognition = 99%

Automatic Speech Recognition and Understanding (ASRU)



Automatic Speech Recognition and Understanding (ASRU)



Use ASRU without modifications for Zurich En-Route





edelweiss four four bravo
climb flight level three three
zero increase mach point seven
five



~~Performance:~~
~~Word Error Rate = 3%~~
~~Command Recognition = 92%~~
~~Callsign Recognition = 99%~~

Automatic Speech Recognition and Understanding (ASRU)



Use ASRU without modifications for Zurich En-Route





edelweiss four four bravo
climb flight level three three
zero increase mach point seven
five



Performance Drop:
Word Error Rate > 12%
Command Recognition < 70%
Callsign Recognition < 90%

Big Drop in Performance!

Speech Recognition and Understanding (ASRU)



How to improve the performance?

"Just" Contact DLR and ask them to fix it → **With another X months of effort**

"Just" hand-over 50 hours of speech from Zurich en-route traffic → **Do not forget to stay GDPR compliant**

"Just" Manually transcribe 10 hours of speech data → **May take 100 hours of expert working time**

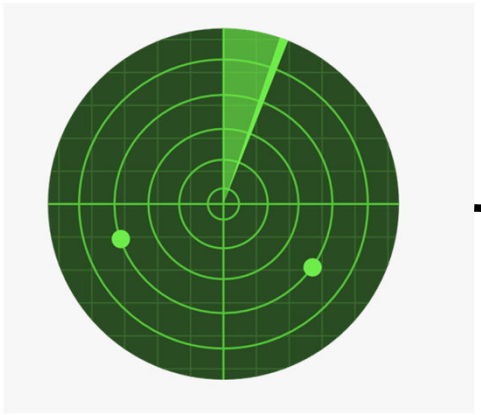
"Just" manually annotate 10 hours of the transcriptions → **Again takes 100 hours of expert working time**

Adaptation and maintenance costs quickly become higher than license costs!

Source: SESAR W2-96 ASR-TLR6 Final CBAT
https://www.sesarju.eu/sites/default/files/documents/solution/SolPj.10-W2-96%20ASR%20CBA%20TRL6.pdf


SELF-MADE-ATC

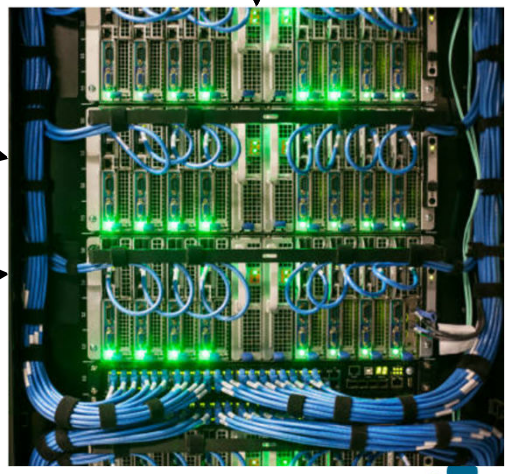
Concept SELF-MADE-ATC



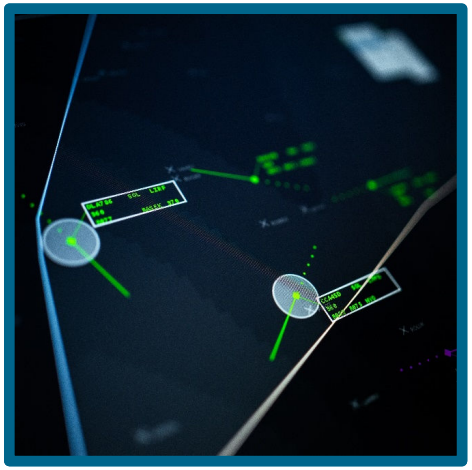


AIRAC

AIP

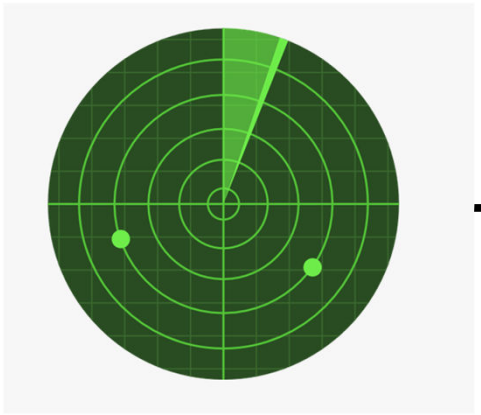


SELF-MADE-ATC

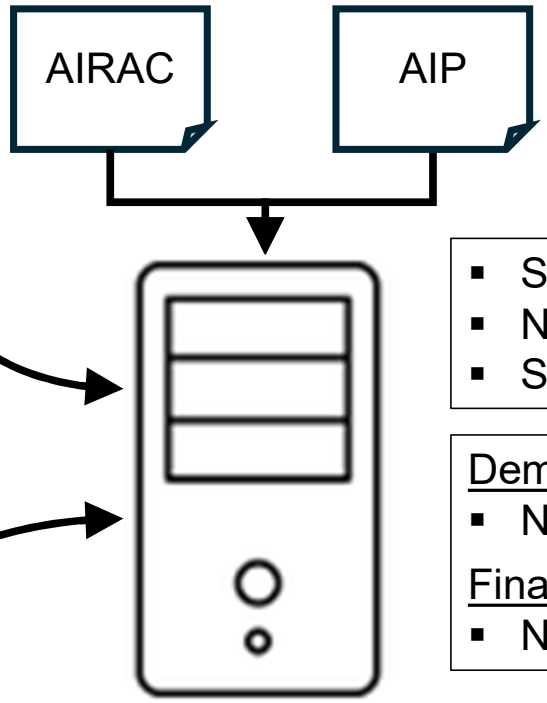


Callsign Highlighting

Concept SELF-MADE-ATC







- SELF-MADE-ATC runs on a single PC
- No internet access required
- Software and models are locally hosted

Demo-PC runs on:
- NVIDIA RTX 6000 with 48 GB VRAM

Final System aims at:
- NVIDIA RTX 5090 with 32 GB VRAM

SELF-MADE-ATC

Features of SELF-MADE-ATC

Self-Maintain → Automatically **maintain** ASRU performance also in case of changes

Self-Adapt → Automatically **adapt** ASRU to new environments like airspace or airports

Self-Deploy → Automatically **deploy** improved/new ASRU versions when available

Self-Evaluate → Automatically **evaluate** and monitor ASRU performance at all time

ASRU = Automatic Speech Recognition and Understanding

Functions:

- **DATA-COLLECTION:** Collects and processes all required data elements automatically
- **MODEL-TRAINING:** Takes gathered data to automatically update/modify required models
- **ADAPT/MAINTAIN:** Automatic and tool-supported adaptation/maintenance for all environments
- **EVALUATE/MONITOR:** Evaluation and Monitoring of all relevant performance metrics at all times

→ **Functions can be executed at the user facilities – No data sharing with third parties!**

Benefits of SELF-MADE-ATC



How to improve the performance?

"Just" Contact DLR and ask them to fix it → **With another X months of effort**

"Just" hand-over 50 hours of speech from Zurich en-route traffic → **Do not forget to stay GDPR compliant**

"Just" Manually transcribe 10 hours of speech data → **Easily takes 100 hours of expert working time**

"Just" manually annotate 10 hours of the transcriptions → **Again takes 100 hours of expert working time**

SELF-MADE-ATC

Benefits of SELF-MADE-ATC



How to improve the performance?

~~"Just" Contact DLR and ask them to fix it → **With another X months of effort**~~
~~"Just" hand-over 50 hours of speech from Zurich en-route traffic → **Do not forget to stay GDPR compliant**~~
~~"Just" Manually transcribe 10 hours of speech data → **Easily takes 100 hours of expert working time**~~
~~"Just" manually annotate 10 hours of the transcriptions → **Again takes 100 hours of expert working time**~~

↓

SELF-MADE-ATC

↓

"No" need of DLR support (just as backup) → **Just need one month of adaptation effort**
"No" data sharing with third parties data stays within the ANSP/airport facilities → **No GDPR trouble**
"No" manual transcription (only for evaluation) → **Only minimum amount of expert working time**
"No" manual annotation (only for evaluation) → **Only minimum amount of expert working time**



1. Motivation for SELF-MADE-ATC
2. **Speech Recognition and Speech Understanding**
3. Building Blocks
   a. Data Collection,
   b. Model Train,
   c. Adapt,
   d. Maintain,
   e. Evaluation,
   f. Monitor
4. Conclusions
5. Overview of today and tomorrow

SELF-MAINTAINING DIGITAL EVOLUTION FOR AIR TRAFFIC CONTROL SUPPORT

8:50

Automatic Speech Recognition and Speech Understanding

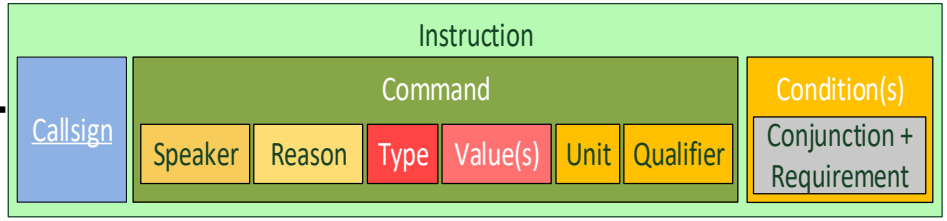
Automatic Speech Recognition alone is not enough.

Speech Recognition is NOT Speech Understanding
Alan Turing 1952

good evening alpine two mike romeo reaching flight level two four zero on course to resia request flight level tw six zero

EJU2MR GREETING
EJU2MR REPORTING ALTITUDE 240 FL
EJU2MR REPORTING DIRECT_TO RESIA
EJU2MR REQUEST ALTITUDE 260 FL

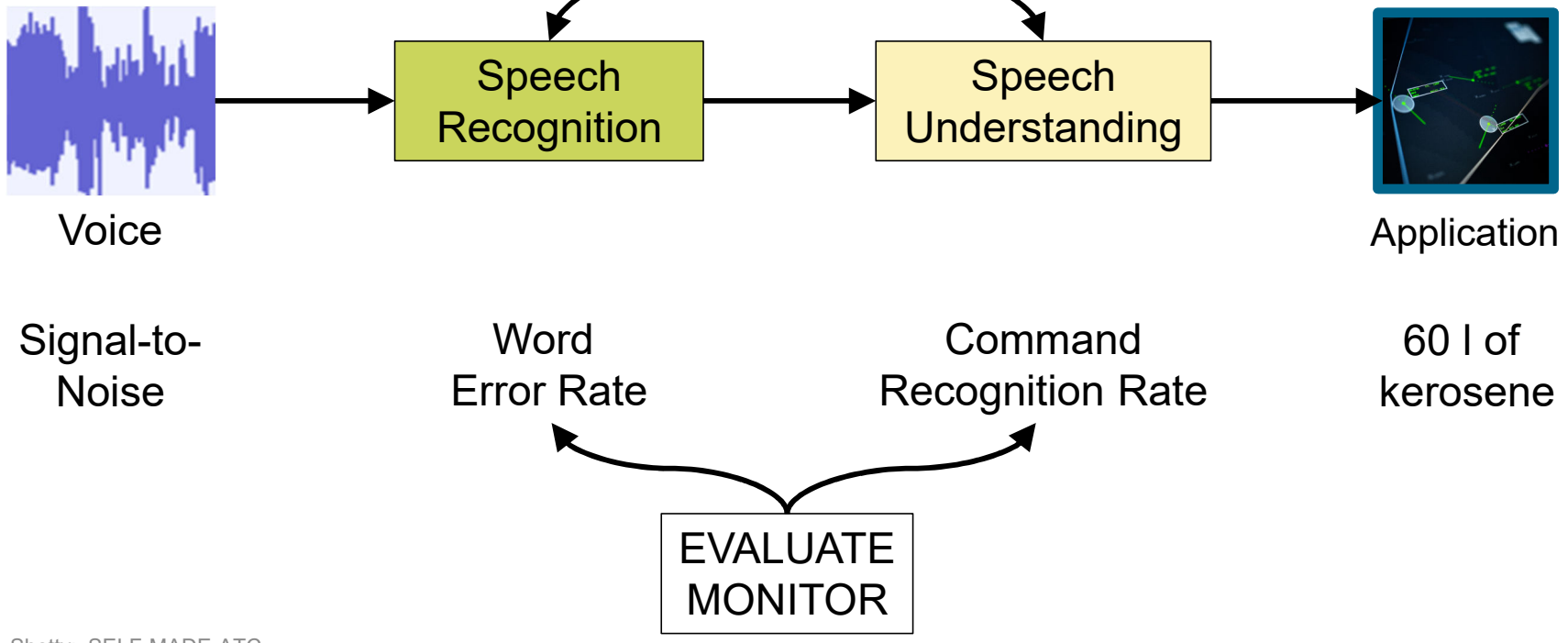
Handed over to EUROCAE as a contribution to standardization the output of ASRU systems as part of EUROCAE Working Group 126



*agreed 2017 between >20 European partners:
Helmke, H., Slotty, M., Poiger, M., Herrer, D.F., Ohneiser, O., Vink, N., Cerna, A., Hartikainen, P., Josefsson, B., Langr, D., García Lasheras, R., Marin, G., Mevatne, O.G., Moos, S., Nilsson, M.N., Pérez, M.B., "**Ontology for transcription of ATC speech commands of SESAR 2020 solution PJ.16-04**," IEEE/AIAA 37th Digital Avionics Systems Conference (**DASC**), London, United Kingdom, 2018

From Voice Signals to Business

SELF-MADE-ATC



Helmke, Kleinert, Feier, Shetty: SELF-MADE-ATC



1. Motivation for SELF-MADE-ATC
2. Speech Recognition and Speech Understanding
3. **Building Blocks**
   a. **Data Collection,**
   b. Model Train,
   c. Adapt,
   d. Maintain,
   e. Evaluation,
   f. Monitor
4. Conclusions
5. Overview of today and tomorrow

SELF-MAINTAINING DIGITAL EVOLUTION FOR AIR TRAFFIC CONTROL SUPPORT

10:00

Data Collection – What's the Input?

Core Input: Unlabelled Audio **(Required)**

- Pure ATCO-pilot voice transmissions
- **No human labelling effort**, just recording effort

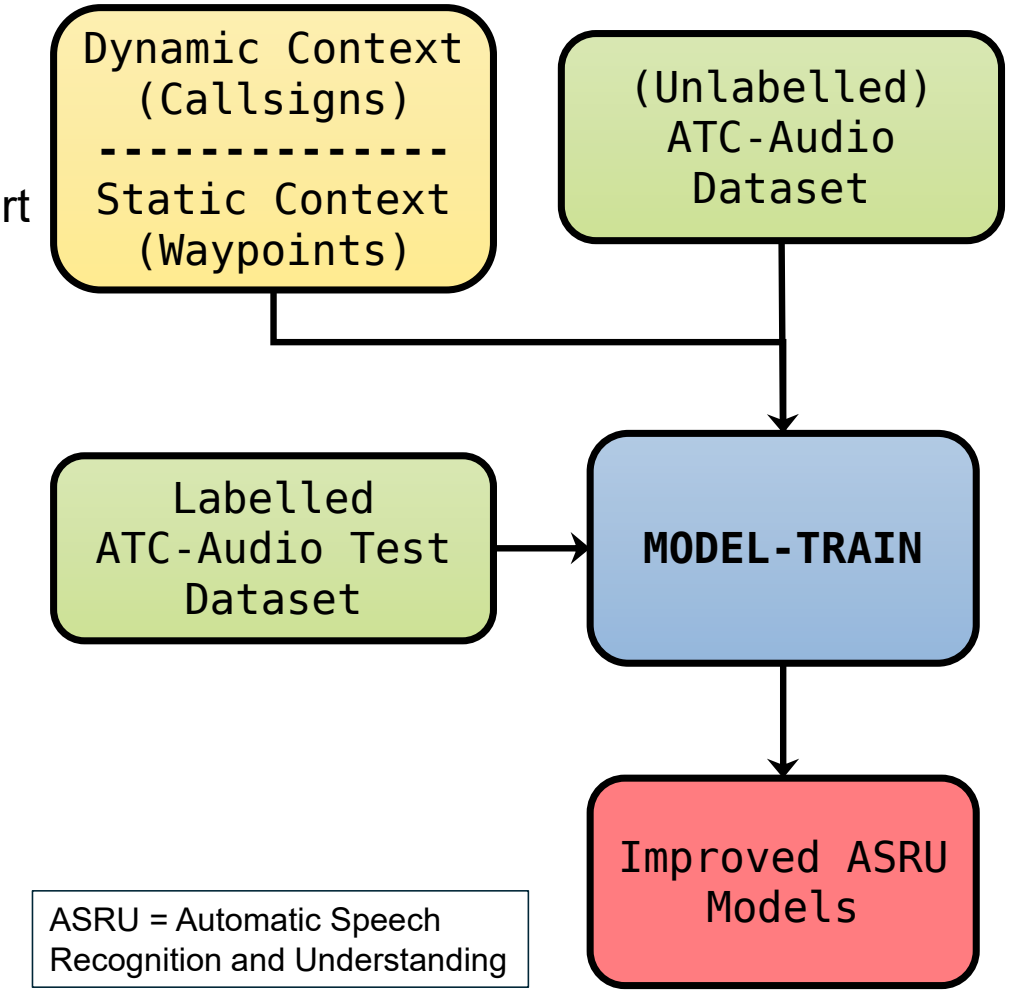
Verification: Labelled Test Set **(Required)**

- Only minimal expert-labelled test set
- Strictly used for performance benchmarking

Booster: Contextual Data **(Optional)**

- Static → Airspace structure, e.g., waypoints
- Dynamic → Active airspace callsigns

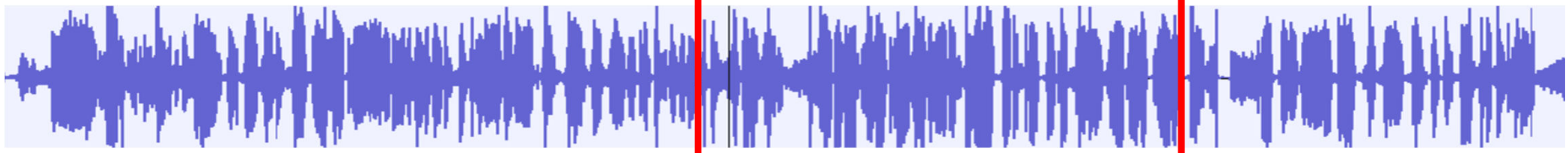
Output→ Ready-to-deploy ASRU models aligned to the target airspace



ASRU = Automatic Speech Recognition and Understanding

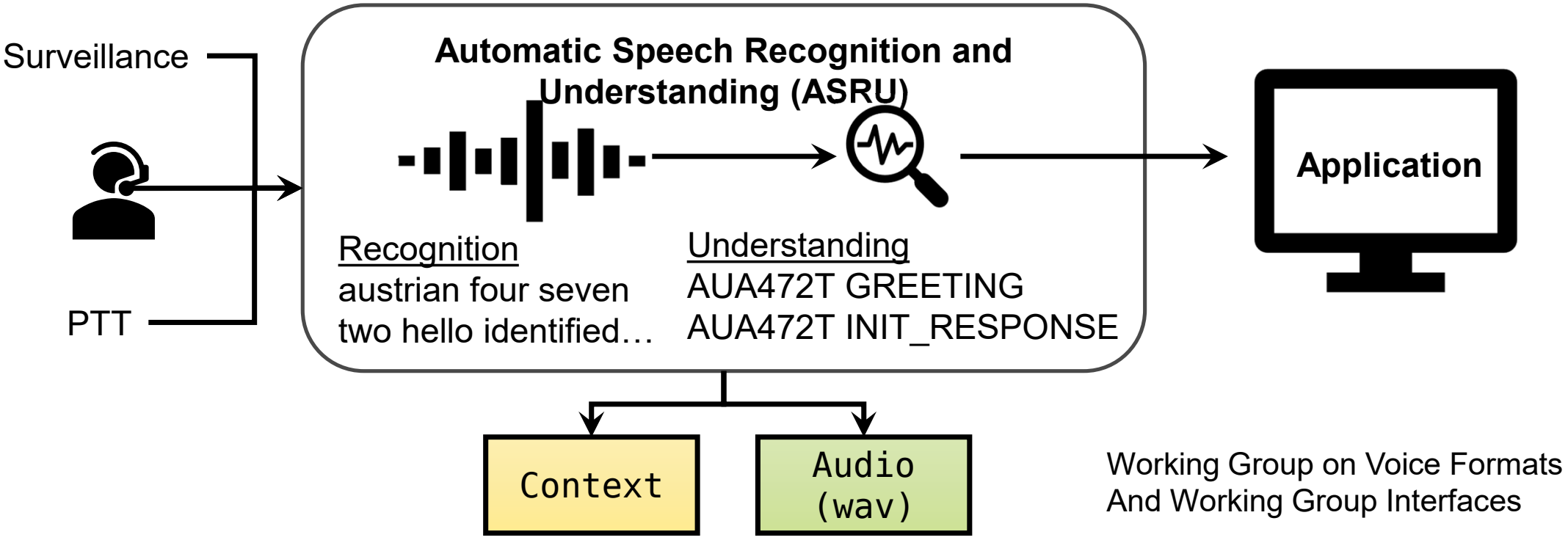
Data Collection – Audio and Transcripts

- **ASR Target:** 16 kHz linear PCM mono (baseline for modern ASR models) – Standard 8 kHz encoded VCS audio fully supported through resampling/transcoding
- **Individual Files:** One file per transmission with speaker role (ATCO or Pilot) – Obtainable through existing PTT signaling and squelch
- **Performance Benchmark:** Only for the small amount of verification data
  - Transcripts – Manual checked word-by-word content of the audio file
  - Annotations – Semantic Interpretation of the transcript



Pilot	ATCo	Pilot
good evening alpine two mike romeo reaching flight level two four zero on course to resia request flight level tw six zero	alpine nine two mike romeo hello identified that is copied after resia direct morok	after resia direct to morok alpine nine two mike romeo

Data Collection at Skyguide



Working Group on Voice Formats
And Working Group Interfaces

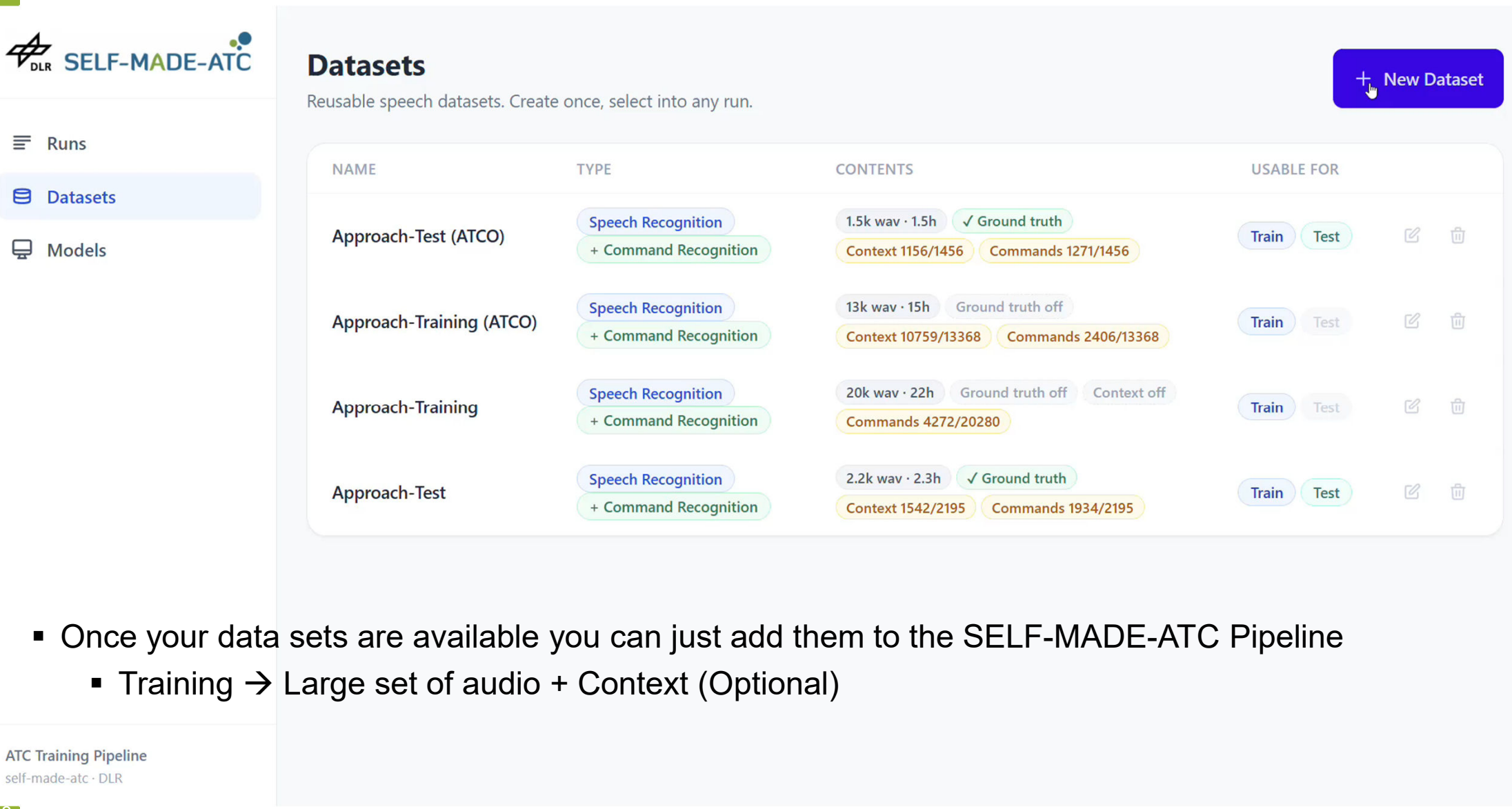
ASRU integrated in ops room as in Zurich already collects all the needed data.



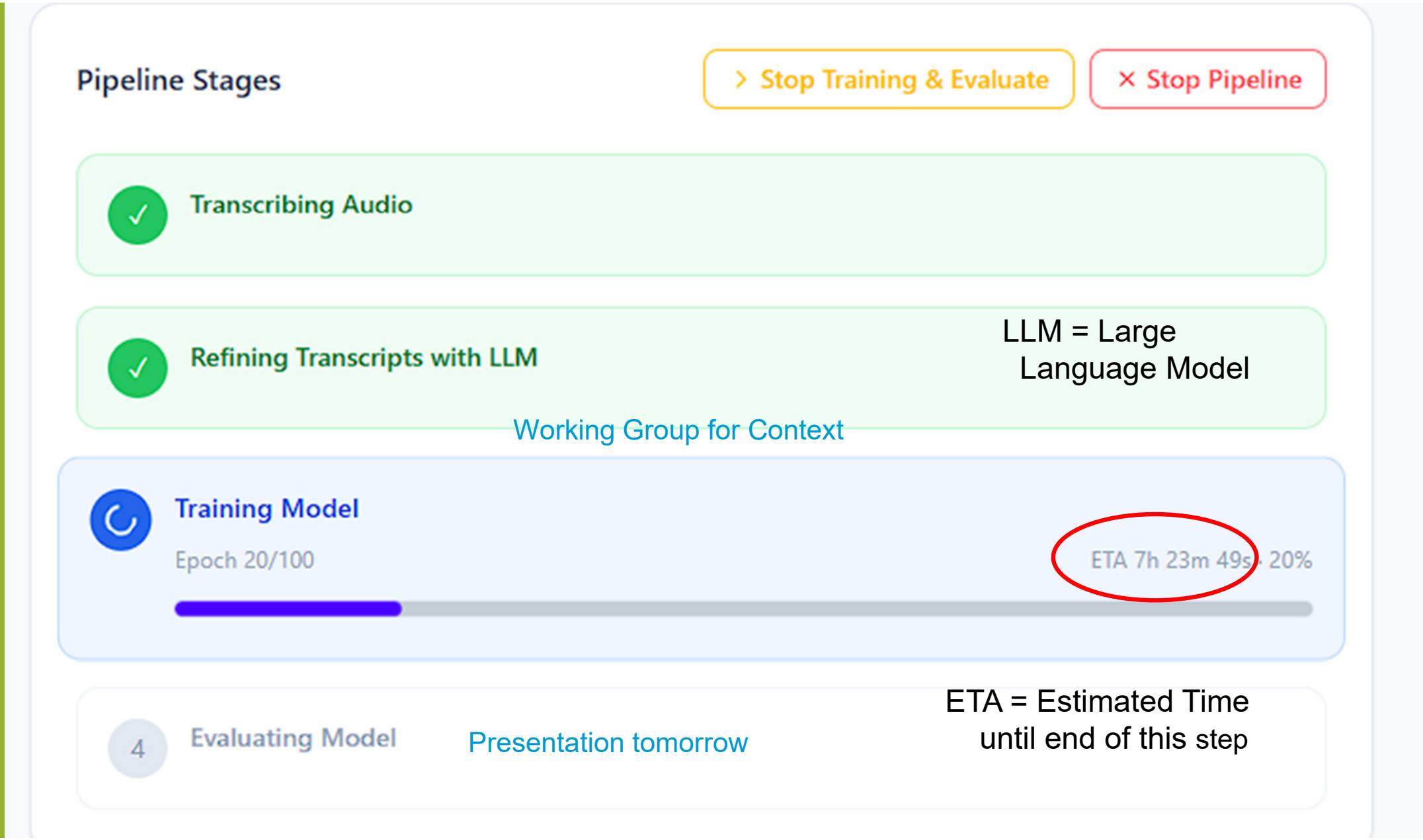
1. Motivation for SELF-MADE-ATC
2. Speech Recognition and Speech Understanding
3. Building Blocks
   a. Data Collection,
   b. Model Train,
   c. Adapt,
   d. Maintain,
   e. Evaluation,
   f. Monitor
4. Conclusions
5. Overview of today and tomorrow

SELF-MAINTAINING DIGITAL EVOLUTION FOR AIR TRAFFIC CONTROL SUPPORT

12:45



- Once your data sets are available you can just add them to the SELF-MADE-ATC Pipeline
  - Training → Large set of audio + Context (Optional)


Pipeline Stages
Stop Training & Evaluate
Stop Pipeline
Transcribing Audio
Refining Transcripts with LLM
LLM = Large Language Model
Working Group for Context
Training Model
Epoch 20/100
ETA 7h 23m 49s · 20%
4
Evaluating Model
Presentation tomorrow
ETA = Estimated Time until end of this step

Model Training – A few Numbers

Approx. 21.7 h training

Step	Runtime
Overall	10h 56m

2.3 h of test data

Data Source	Data Size	WER ATCo	Length [h]
Baseline	0	14.2%	---
Automatic Transcriptions	20280	7.2%	21.7

WER = Word Error Rate
Data Size = number of wav files

Csgn Context	WER ATCo
no	8.7%
yes	7.2%

Model Training – A few Numbers (2)

Data Source	Data Size	WER ATCo	Length [h]
Baseline	0	14.2%	---
Manual	4476	7.1%	4.0
Automatic	20280	7.2%	21.7

WER = Word Error Rate
Data Size = number of wav files

1 hour of transcription costs 10-20 hours of expert work
1 hour of annotation costs the same effort again.
With SELF-MADE-ATC you need “only” run time.
For the 4 hours you save roughly 40-120 hours of expert work.

4 hours:	60 hours
40 hours:	600 hours
400 hours:	6000 hours

You save it for each airport.
You save it for each additional modification.



1. Motivation for SELF-MADE-ATC
2. Speech Recognition and Speech Understanding
3. Building Blocks
   a. Data Collection,
   b. Model Train,
   c. Adapt,
   d. Maintain,
   e. Evaluation,
   f. Monitor
4. Conclusions
5. Overview of today and tomorrow

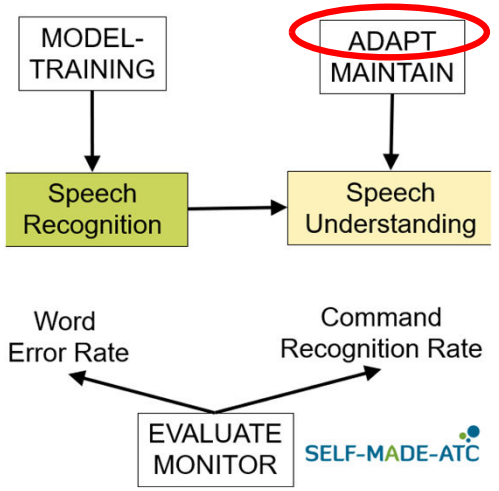
SELF-MAINTAINING DIGITAL EVOLUTION FOR AIR TRAFFIC CONTROL SUPPORT

18:20

Modules ADAPT and MAINTAIN

swiss good evening alpine nine three mike romeo
reaching flight level three four zero
on course to resia
requesting flight level three six zero

EJU93MR STATION SWISS
EJU93MR GREETING
EJU93MR REPORTING ALTITUDE 340 FL
EJU93MR REPORTING DIRECT_TO RESIA none
EJU93MR REQUEST ALTITUDE 360 FL



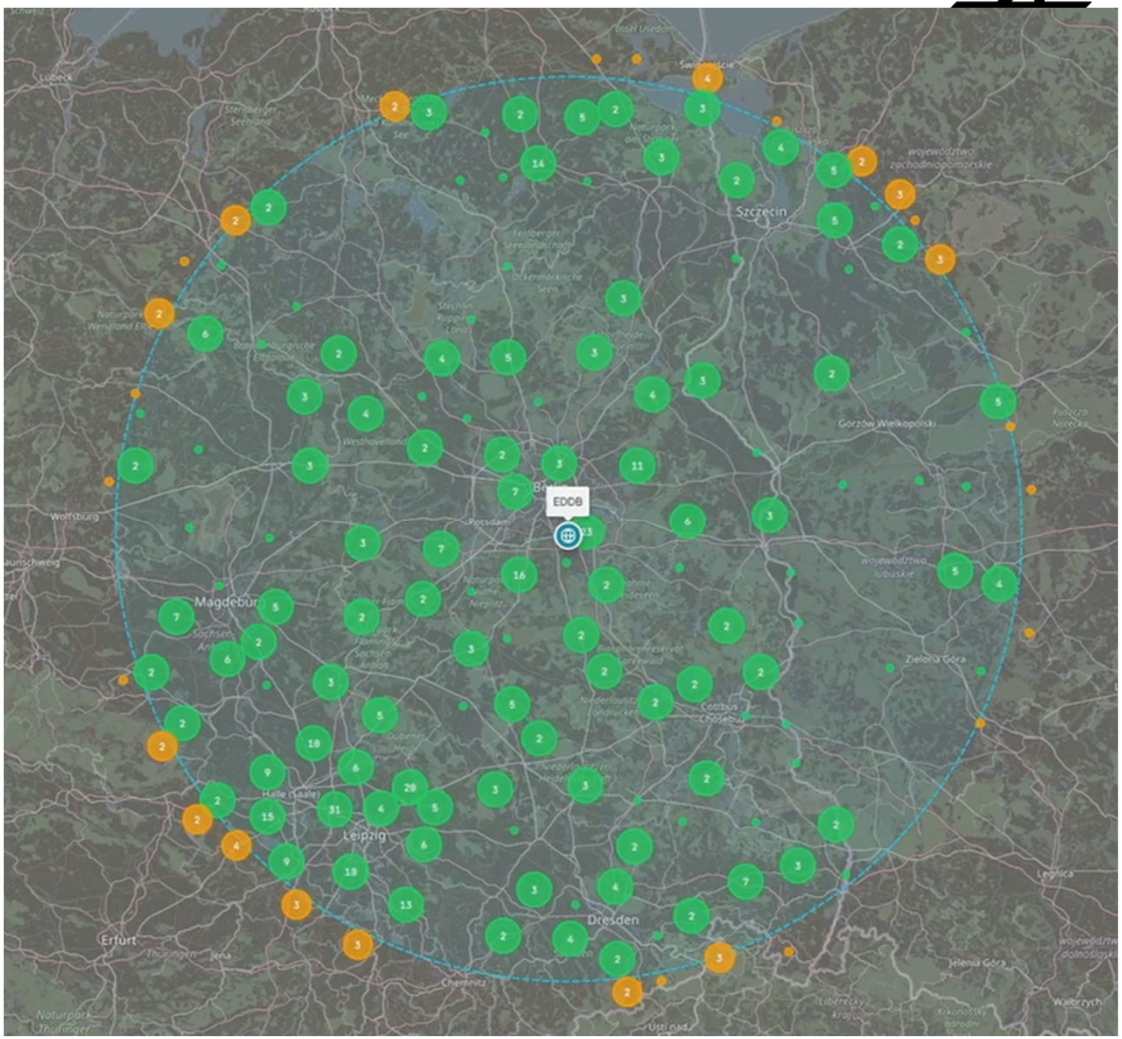
ADAPT and MAINTAIN support Understanding part

Module ADAPT and MAINTAIN

- Update of ATC Concepts
  - New waypoint names like ”resia”
  - Runway names “25L” or “25-Center”
  - Changing runway names “09” to “08” if magnetic changes
  - What is the ATCo saying “two five left”, “two five middle”?
  - Frequency names “swiss radar” and “one nine six decimal four four”

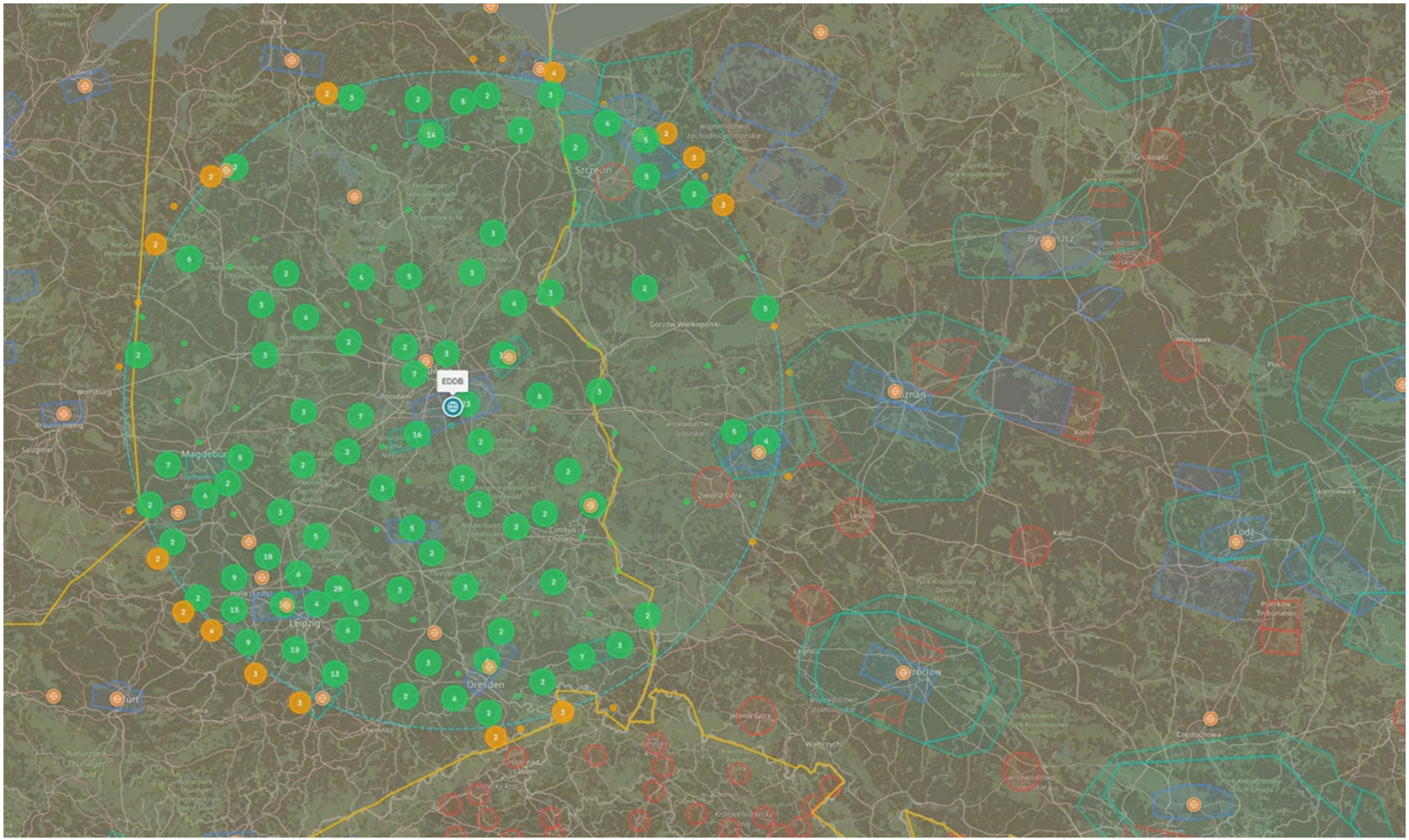
- Phraseology CLIMB: “climb”, “climb to” or “up to”
- Define a new command “START SIMULATOR”
- Adding new airline designator “edelweiss”, “alpine”

Module ADAPT: Adding ATC concepts like waypoints



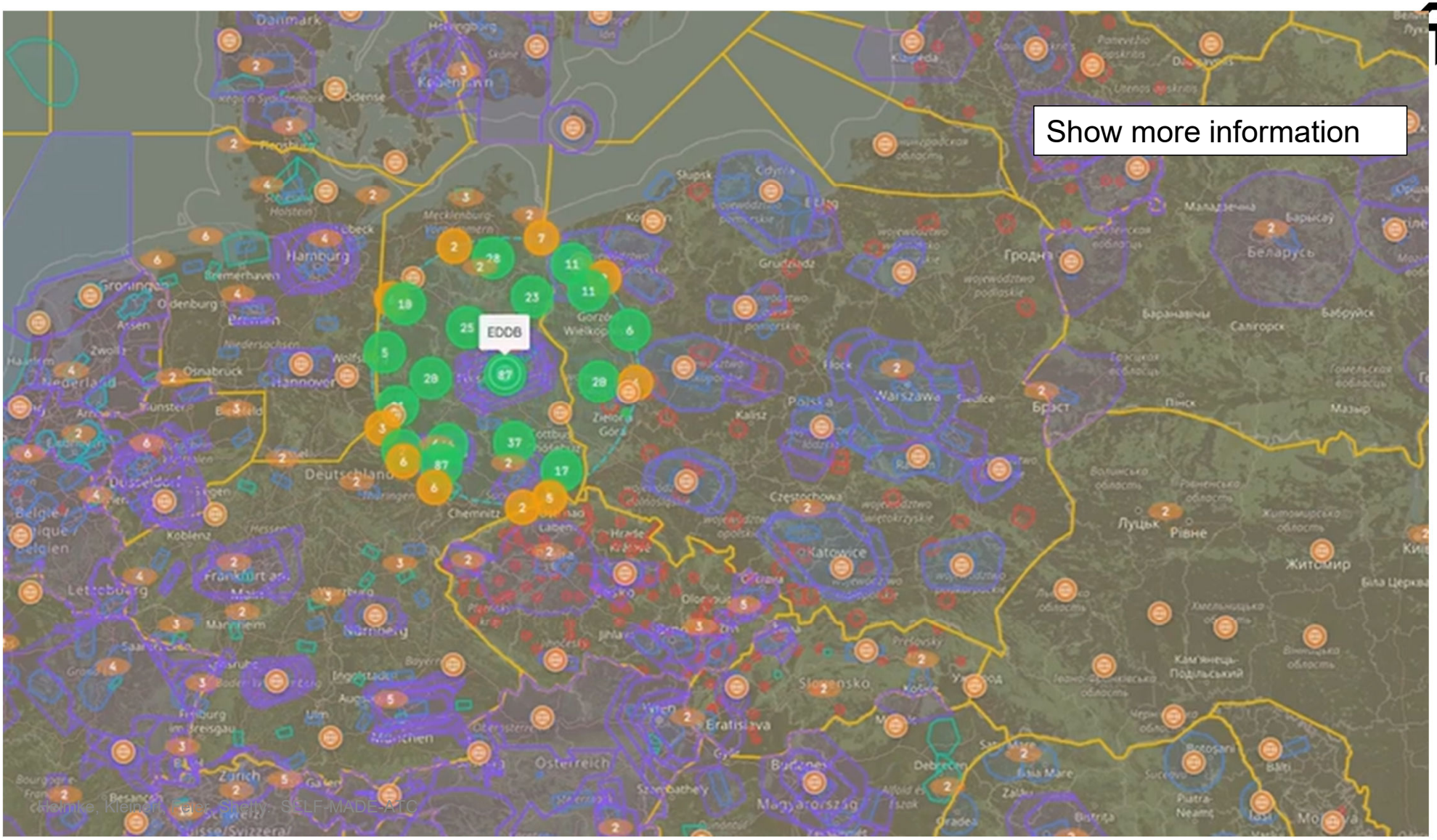
Extraction relevant ATC concepts from external databases and graphically manipulating them.

Module ADAPT: Adding ATC concepts

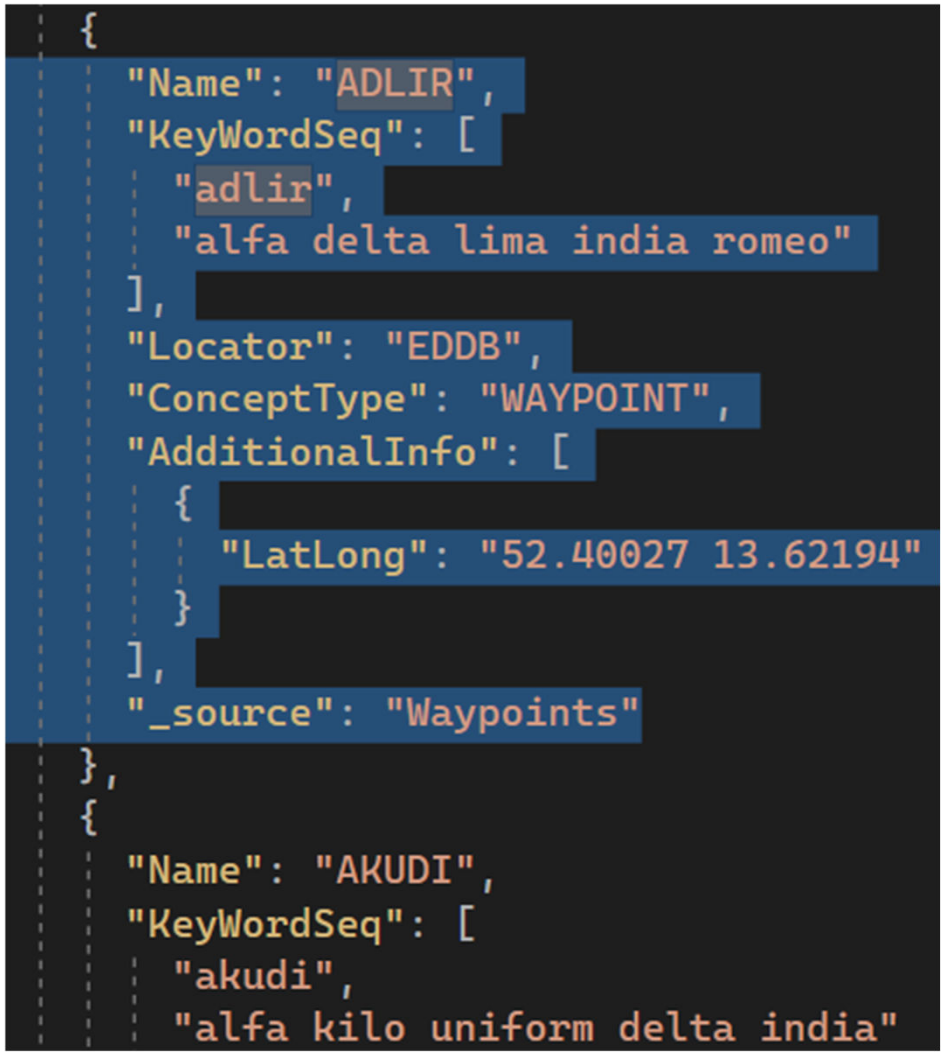


Showing nearby airports.

Module ADAPT: Adding ATC concepts



Module ADAPT: Adding ATC concepts



```
{
  "Name": "ADLIR",
  "KeyWordSeq": [
    "adlir",
    "alfa delta lima india romeo"
  ],
  "Locator": "EDDB",
  "ConceptType": "WAYPOINT",
  "AdditionalInfo": [
    {
      "LatLong": "52.40027 13.62194"
    }
  ],
  "_source": "Waypoints"
},
{
  "Name": "AKUDI",
  "KeyWordSeq": [
    "akudi",
    "alfa kilo uniform delta india"
```

Transformation into JSON files to be graphically manipulated or via simple editors

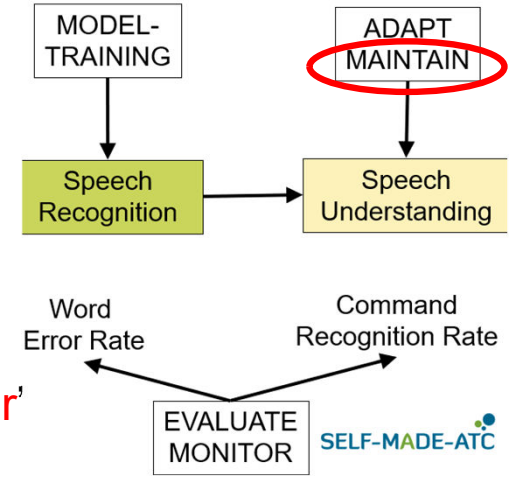
ADAPT is the tool for ATM system suppliers
End-users might use it in simulation environments.

More tomorrow: Presentation of Oliver

Module ADAPT and MAINTAIN

End user organization (MAINTAIN):

- New waypoint names like ”resia”
- Runway names “25L” or “25-Center”
- Changing runway names “09” to “08” if magnetic changes
- What is the ATCo saying “two five left”, “two five middle”?
- Phraseology CLIMB: “climb”, “climb to” or “up to”
- Adding new airline designator “edelweiss”, “alpine”
- Frequency names “swiss radar” and “one nine six decimal four four”
- …



ATM system supplier (ADAPT):

- Define a new command “START SIMULATOR”
- Adding 100 waypoints
- Updating word sequences, if also re-training the recognition models is included
- …

Module MAINTAIN in DLR's Simulators

- Speech Recognition and Understanding used to support simulation pilots
- Today Singapore, tomorrow Maastricht sector, …
- 500 new waypoints, some of them only used
- Training data you have at the end of the validation trials
- Synthetic data created by speech generation is used



1. Motivation for SELF-MADE-ATC
2. Speech Recognition and Speech Understanding
3. Building Blocks
   a. Data Collection,
   b. Model Train,
   c. Adapt,
   d. Maintain,
   e. Evaluation,
   f. Monitor
4. Conclusions
5. Overview of today and tomorrow

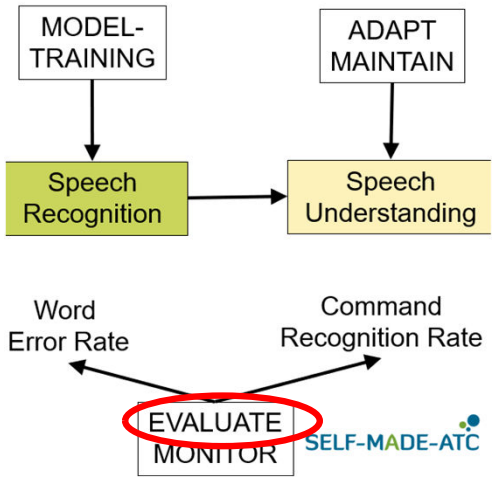
SELF-MAINTAINING DIGITAL EVOLUTION FOR AIR TRAFFIC CONTROL SUPPORT

22:40

Module EVALUATION

- Is the new model better than the currently used one?

- You need voice utterances with manual transcriptions and manual annotations.
- This could be old utterances (ten years old) or the new challenges, which did not work before.
- SELF-MADE-ATC provides tools to calculate different metrics.



Module EVALUATION: Metrics

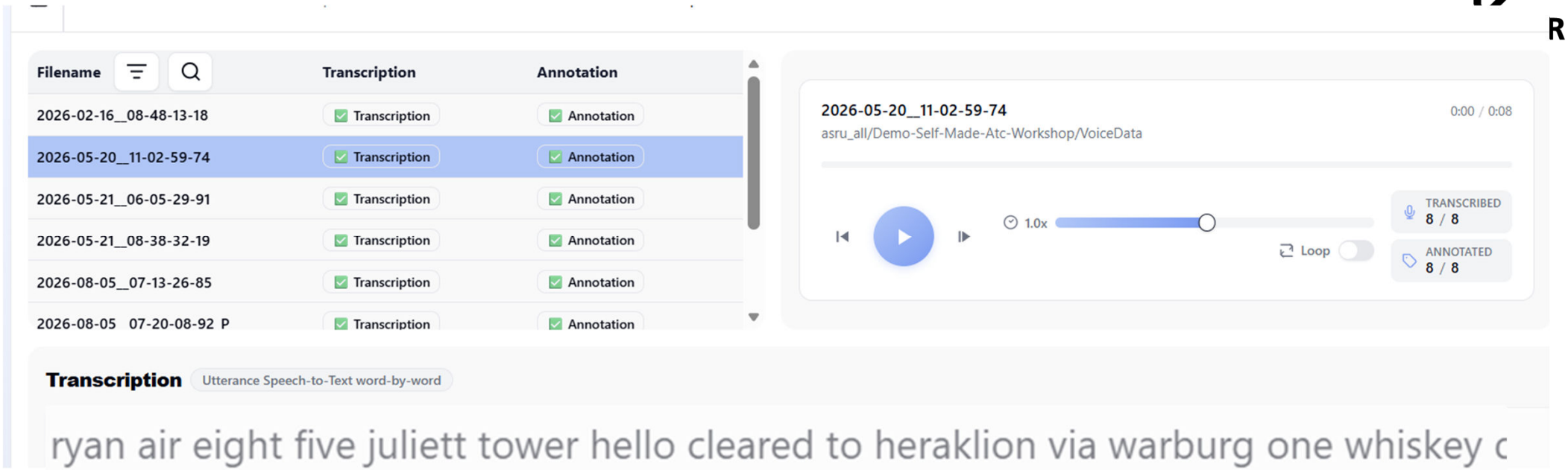


- Word error rates (WER)
- Word error rates for local words like edelweiss or epino or … Top 10 words

- Callsign recognition rate / callsign recognition error rate
- Command recognition (error) rate
- Command recognition (error) rate for e.g. TAXI or DIRECT_TO …

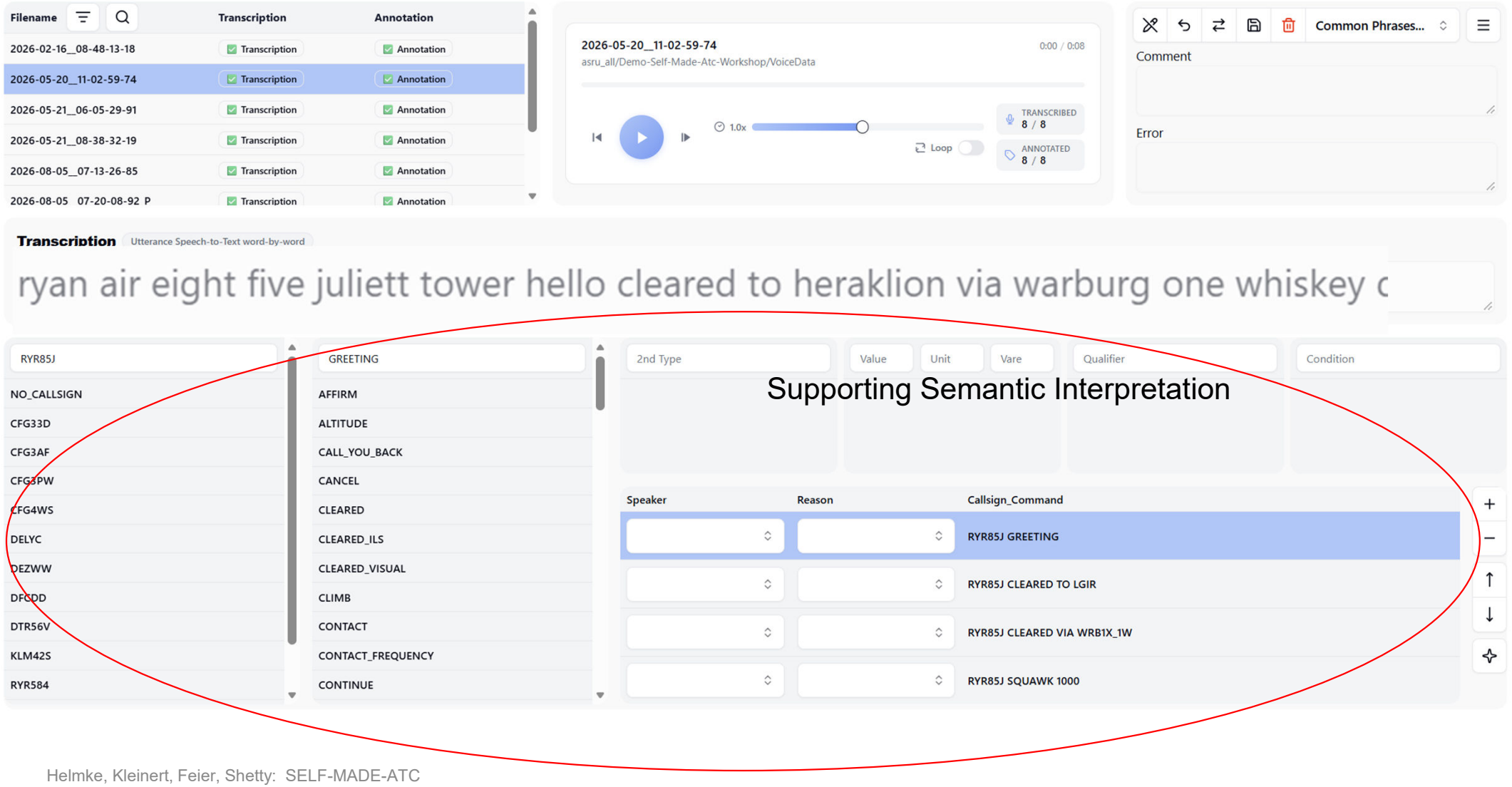
- Metrics for your special application (e.g. readback error false alarm rate, …)

Module EVALUATION: CoCoLoToCoCo

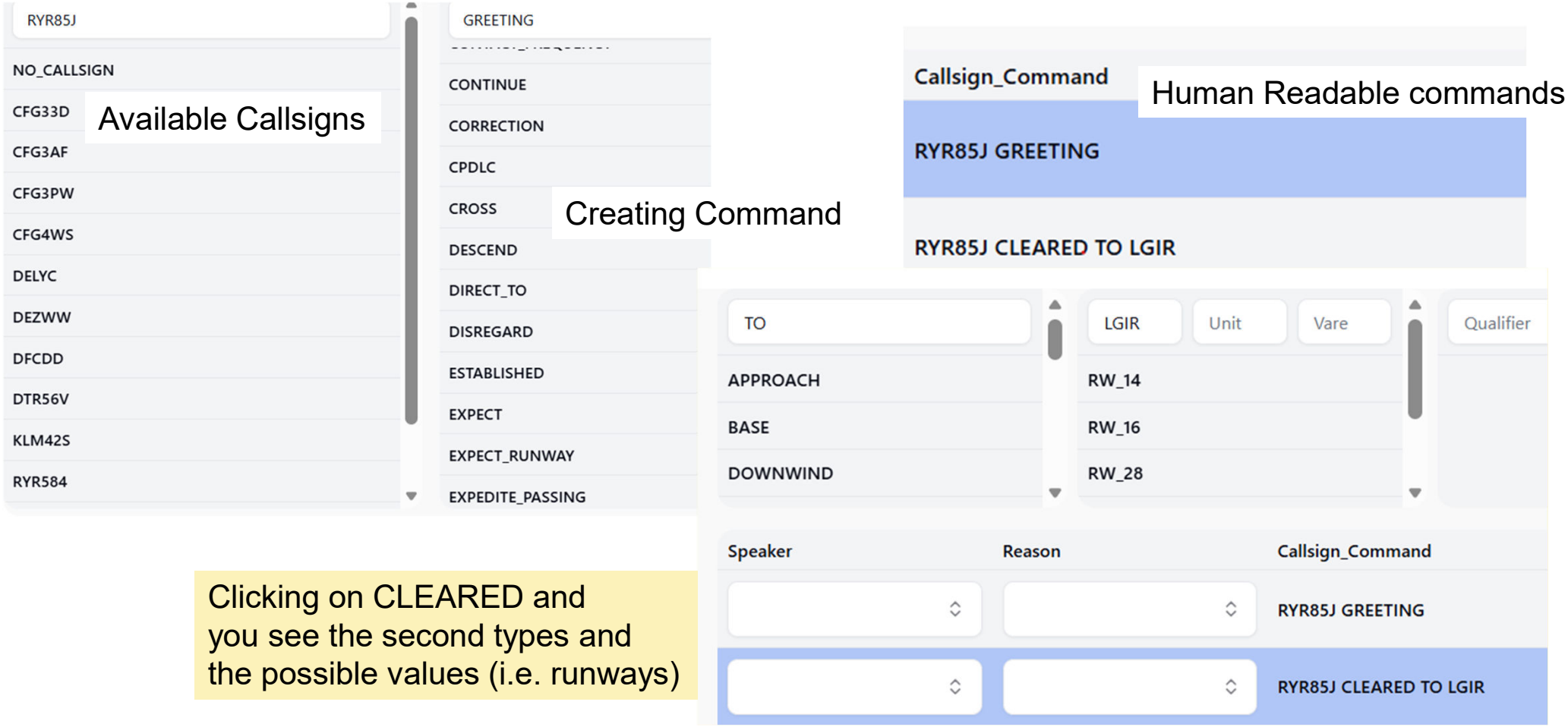


CoCoLoToCoCo is a web-based GUI, which enables to play, transcribe and check wav files.

Module EVALUATION: CoCoLoToCoCo



Module EVALUATION: CoCoLoToCoCo



Clicking on CLEARED and
you see the second types and
the possible values (i.e. runways)



1. Motivation for SELF-MADE-ATC
2. Speech Recognition and Speech Understanding
3. Building Blocks
   a. Data Collection,
   b. Model Train,
   c. Adapt,
   d. Maintain,
   e. Evaluation,
   f. Monitor
4. Conclusions
5. Overview of today and tomorrow

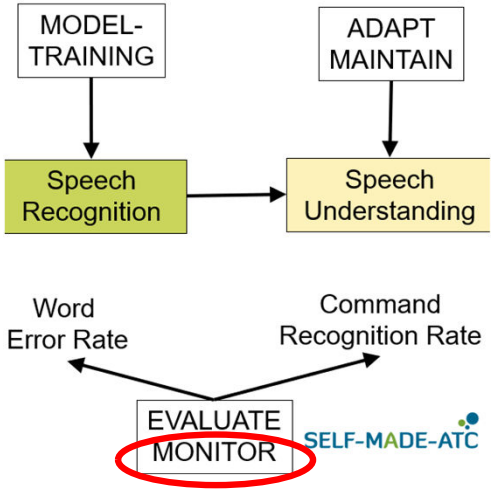
SELF-MAINTAINING DIGITAL EVOLUTION FOR AIR TRAFFIC CONTROL SUPPORT

26:30

Module Monitor

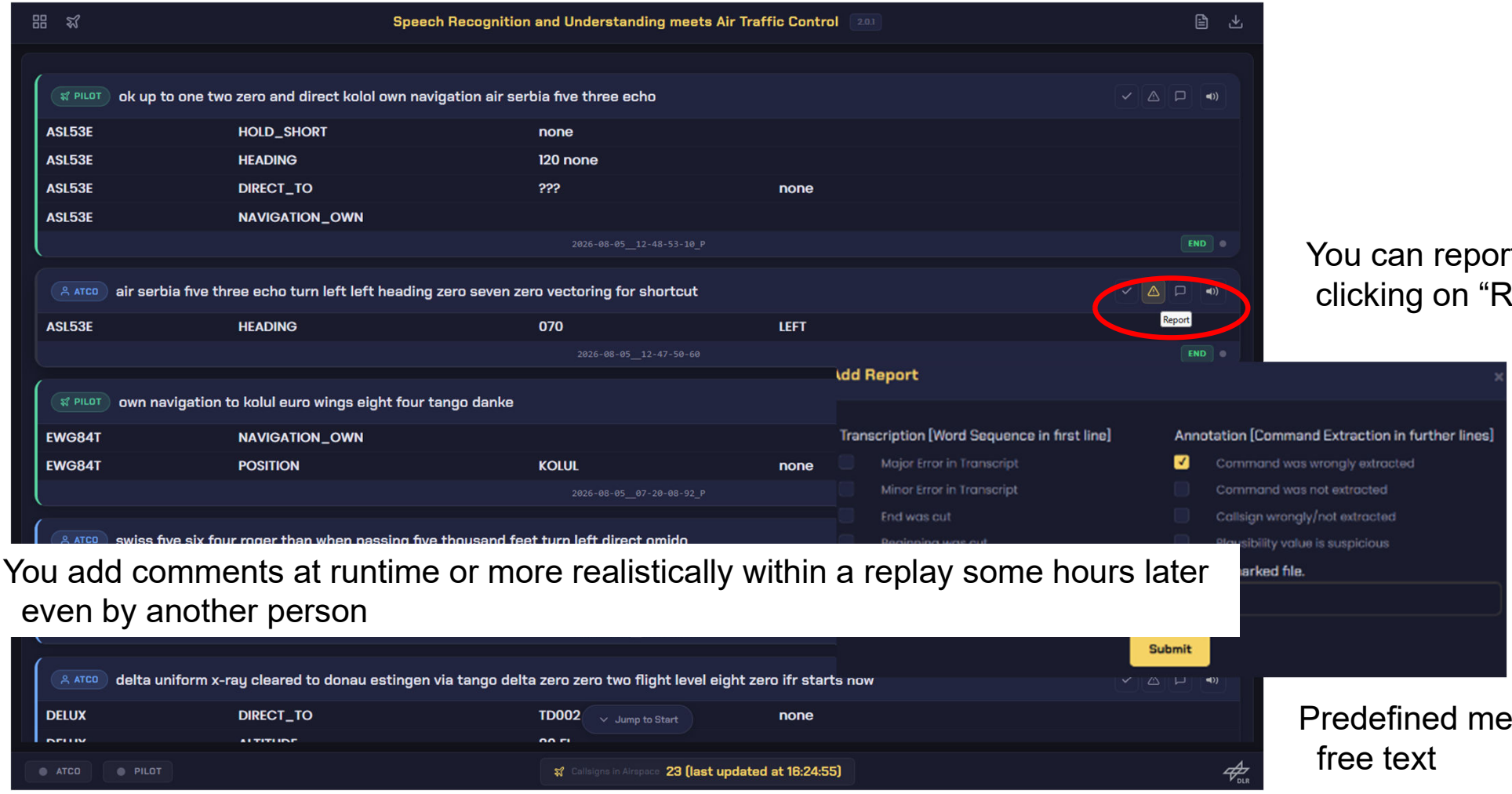
- Testing never proves the absence of errors.
- You (including the end-user) need to monitor at runtime.





Helmke, Kleinert, Feier, Shetty: SELF-MADE-ATC

Module Monitor with SpeechLog



You can report with clicking on “Report”

You add comments at runtime or more realistically within a replay some hours later even by another person

Predefined messages and free text

More in the Working Groups related to Safety

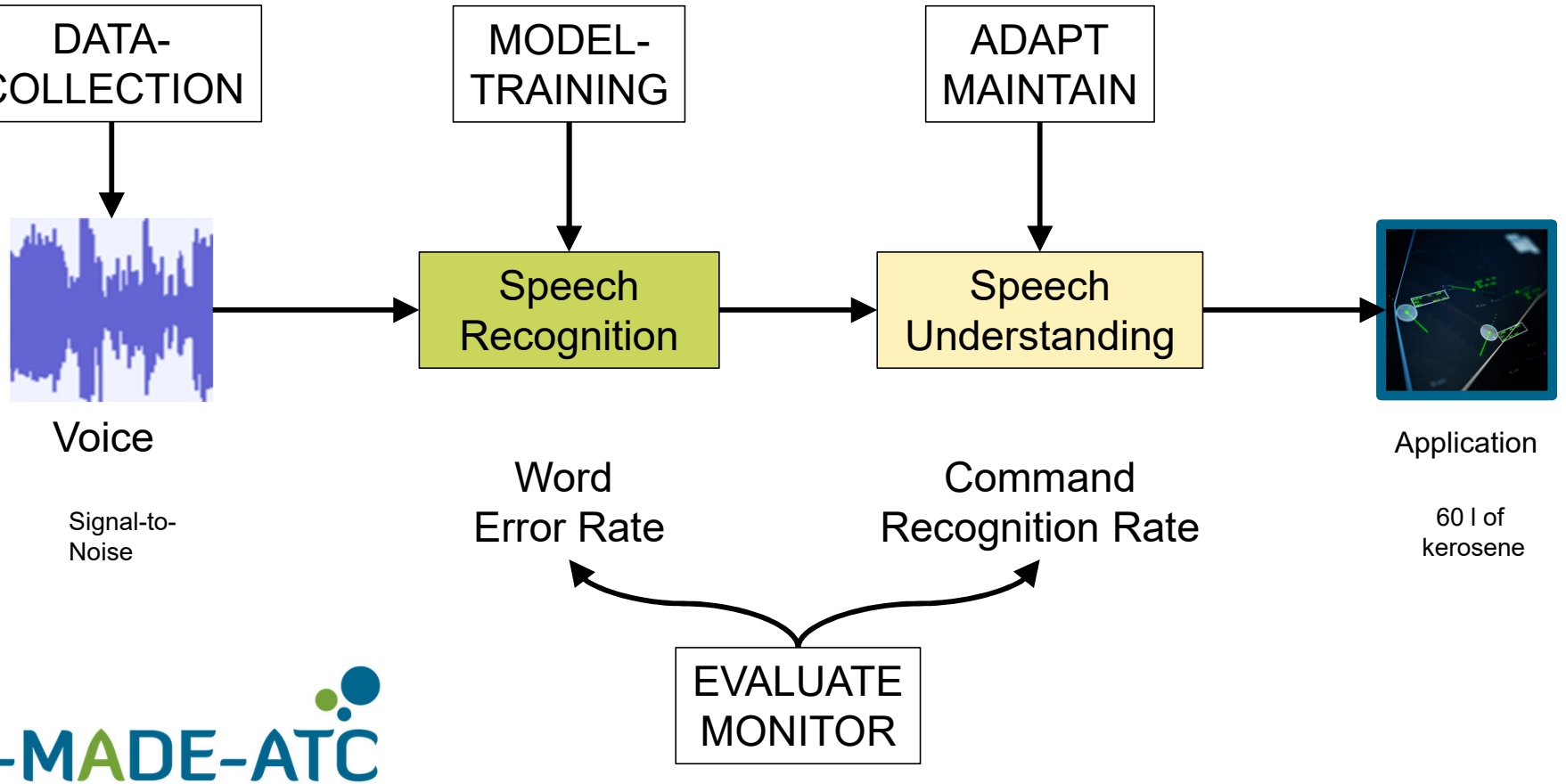


1. Motivation for SELF-MADE-ATC
2. Speech Recognition and Speech Understanding
3. Building Blocks
   a. Data Collection,
   b. Model Train,
   c. Adapt,
   d. Maintain,
   e. Evaluation,
   f. Monitor
4. **Conclusions**
5. Overview of today and tomorrow

SELF-MAINTAINING DIGITAL EVOLUTION FOR AIR TRAFFIC CONTROL SUPPORT

25:45

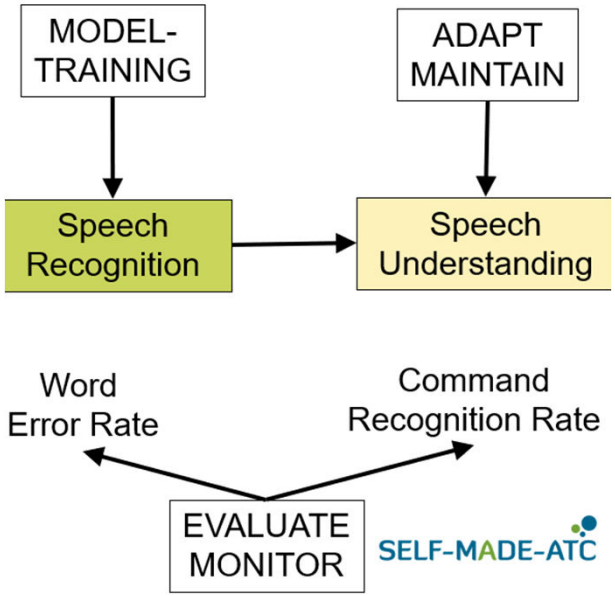
Building Blocks for Adapting ASRU



SELF-MADE-ATC

Conclusions

- Speech Recognition is NOT Speech Understanding.
- Many time-consuming manual steps can be omitted.
- Issues with personal data do not exist. Data stays with the end-users
- We cover the whole chain from the speech signal to the validation of the application plus maintenance and adaptation
- Task sharing between end-user and ATM supplier



And last, but not least:
DLR can offer a test license to end-users and ATM system suppliers.

Impressum

Topic:	SELF-MADE-ATC Self-Maintaining Digital Evolution for Air Traffic Control Support Current Status of Tool Development
Date:	2026-11-04
Authors:	Prof. Dr. Hartmut Helmke, Matthias Kleinert, Niclas Feier, Shruthi Shetty
Institute:	Institute of Flight Guidance (FL)
Credits:	All figures and photos „DLR (CC BY-NC-ND 3.0)“, if not otherwise specified