



Machine learning in speech recognition

Hartmut Helmke, Project Leader, DLR
Aneta Cerna, Project leader, ANS of the Czech Republic
Petr Hlousek, Project leader, ANS of the Czech Republic

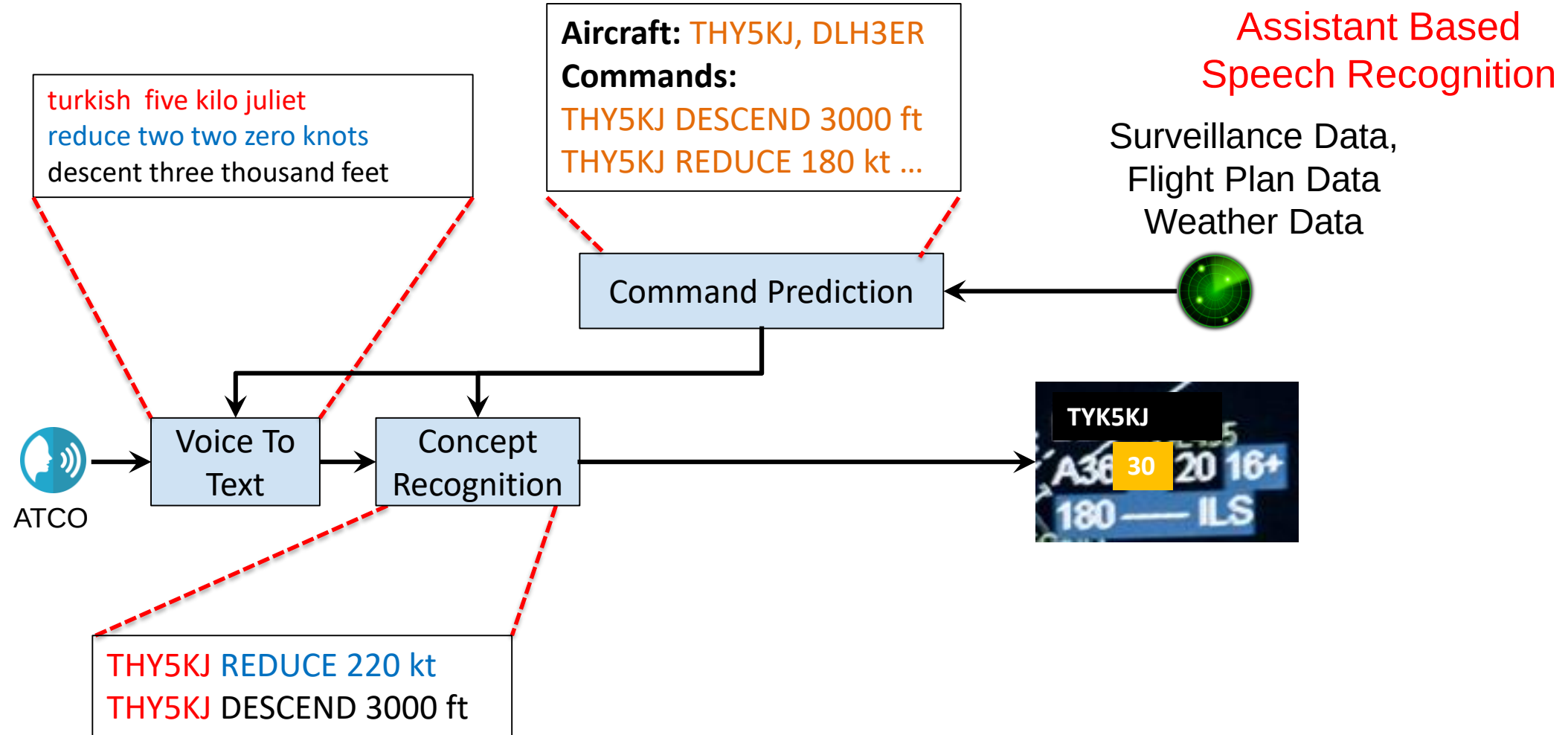


Contents

1. Overview of Assistant Based Speech Recognition in ATM
2. The view of an ANSP:
Results from Trials with
Nuance Recognition Engine



From the Controller Utterance to the HMI



Applications of Speech Recognition in ATM

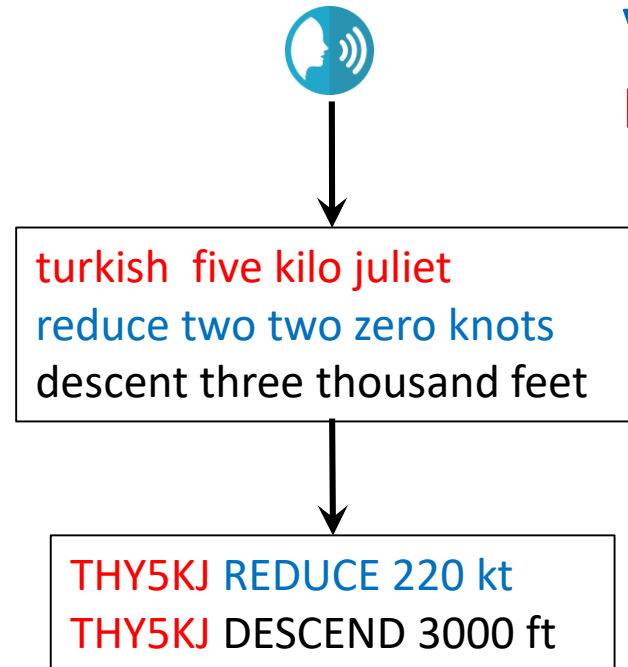


Pre-Filling of Radar Label Entries

1. Callsign Highlighting (Direction Finder Replacement)
2. Keyword Spotting
3. Integration with CPDLC
4. Workload Prediction
5. Read-back Error Detection
6. ...

What is Needed For ASR?

1. Most important are DATA
2. Second: DATA
3. Third: DATA

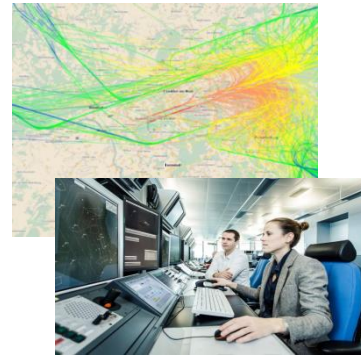


Voice Utterances
Data Privacy, Ethics

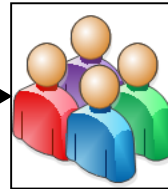
Transcriptions
Expensive

Annotations
Even more expensive
Missing rules

Machine Learning of Speech Recognition Models for Controller Assistance



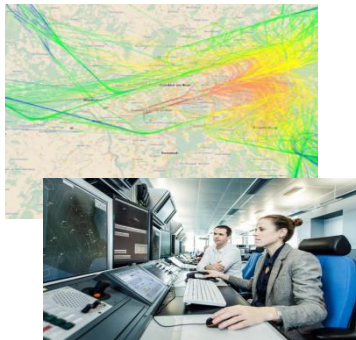
AcListant



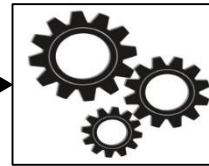
Group of experts

Automatic
Speech
Recognition

Dusseldorf



MALORCA



Automatic Learning
Process

Radar Data

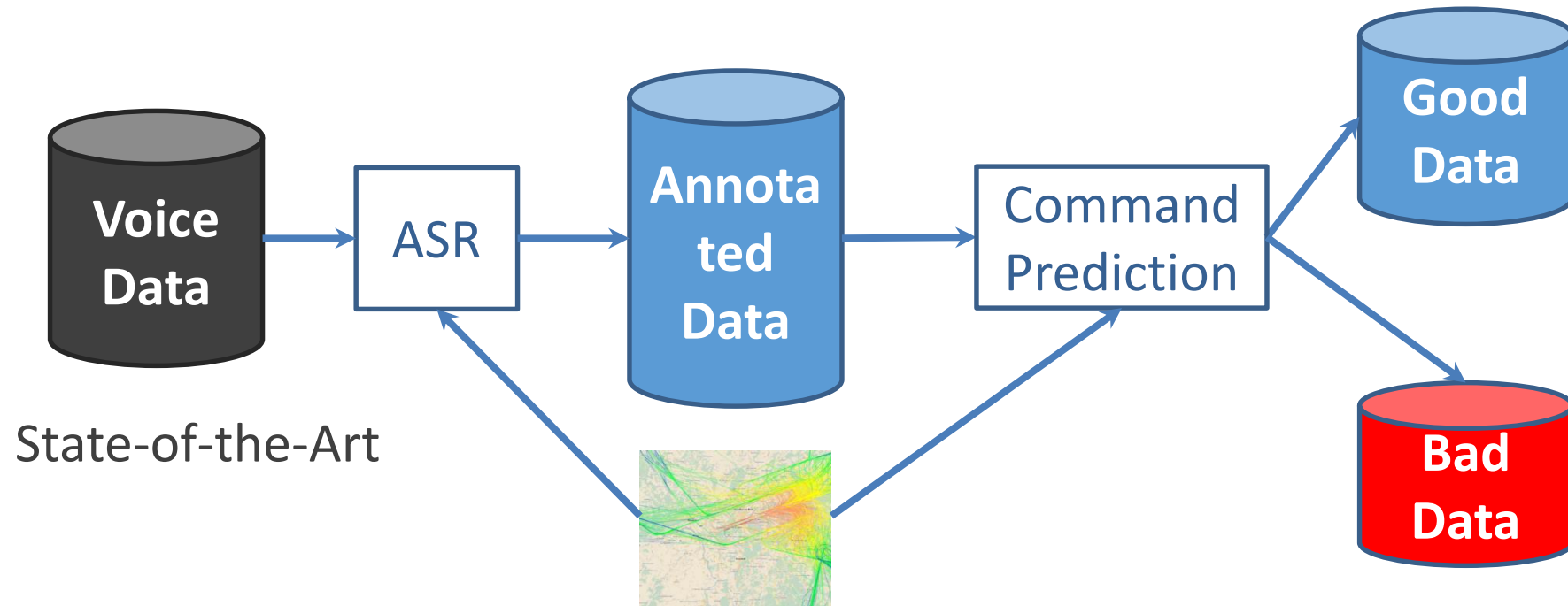
Voice Recordings

Automatic
Based
Speech
Recognition

Prague,
Vienna

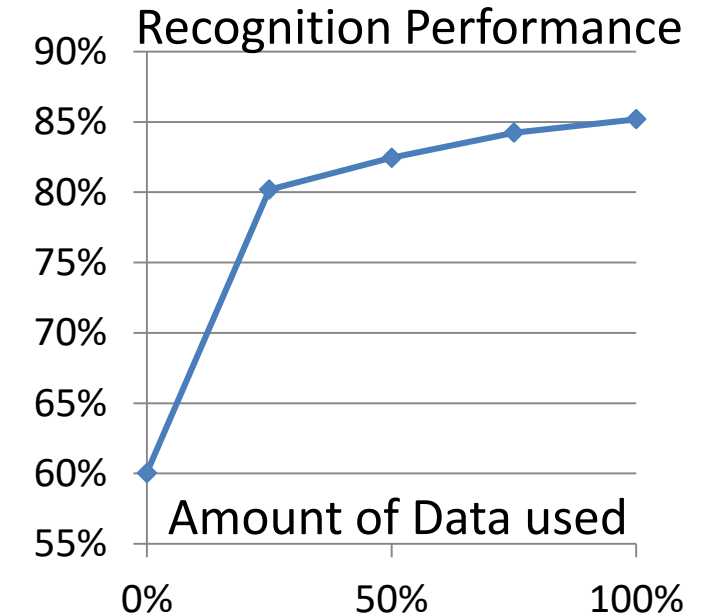
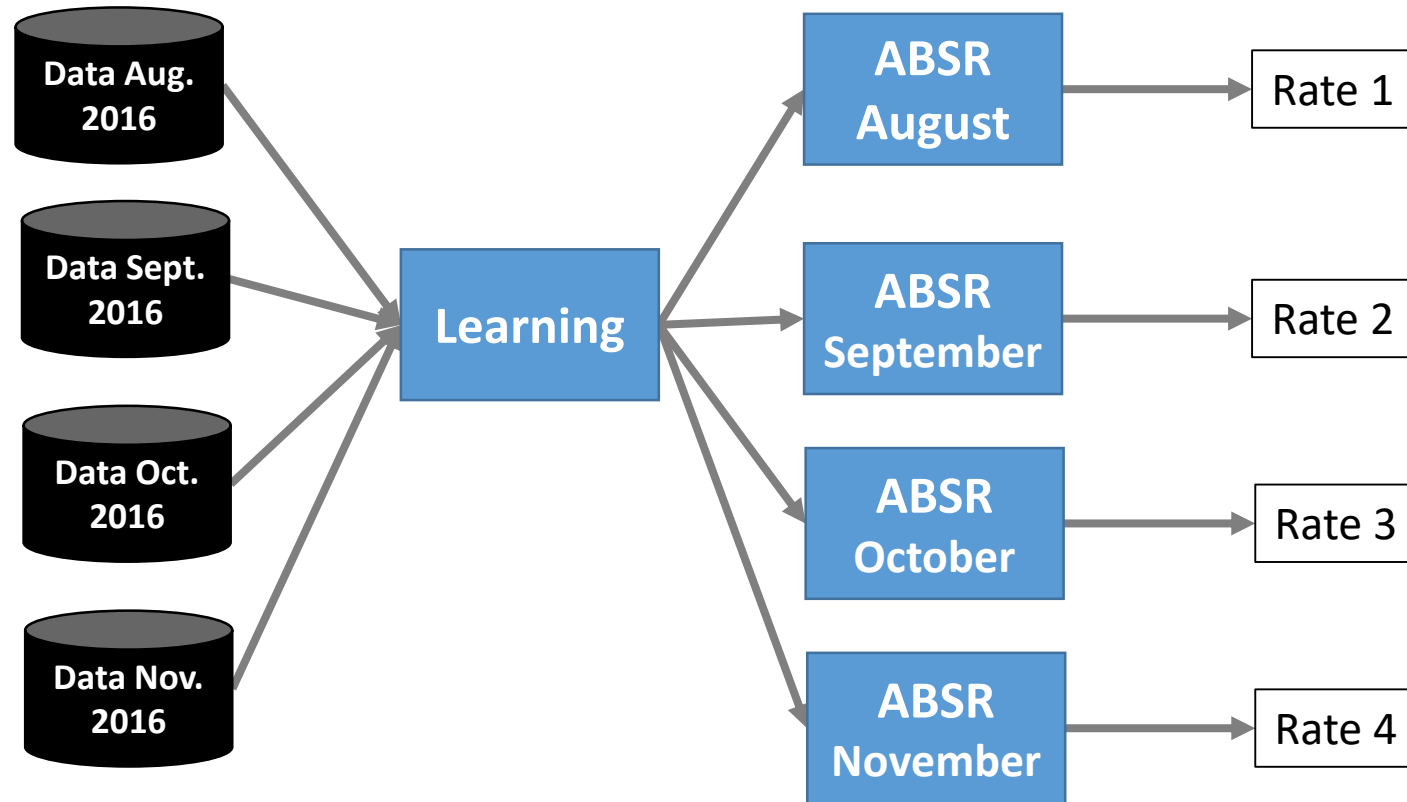
Instead of (highly skilled and paid) **experts**, **machine learning** is used.

Invention of MALORCA



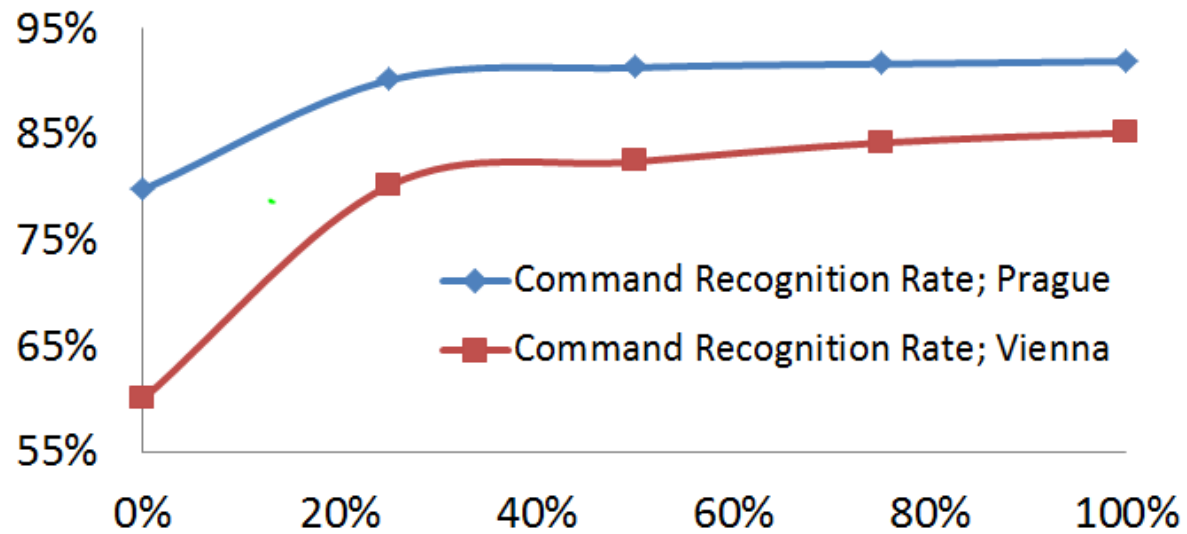
Data selection: Select “good” or “bad” data

MALORCA: Proof-of-Concept



Learning Curve

- Command Recognition depending on amount of provided untranscribed training data



Vienna	Recognition Rate
0%	60.0%
25%	80.2%
50%	82.4%
75%	84.2%
100%	85.2%

Prague	Recognition Rate
0%	79.8%
25%	90.2%
50%	91.3%
75%	91.7%
100%	91.9%

Challenge of MALORCA

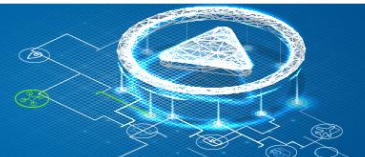


Follow the A320.

- follow the A320
- follow the a tree twenty
- follow the A three twenty
- ...

How to transcribe?

The step from transcriptions
to **annotations** (ATC concepts)
is **ambiguous!!!**



SESAR-WA Solution 16-4

22 partners from 15 European countries agreed on common transcription and annotation rules (**ontology**).

This eases exchange of learning data and comparison of different ASR engines.

Currently DLR developed a tool for annotation.

Automatic learning of the rules.



Annotation/Transcription Ontology CWP HMI

speedbird eight six one resume own navigation proceed direct rapet

- How to annotate (to fix the semantics)?

1. **BAW861** DIRECT TO WP RAPET
2. **BAW861** RAPET NAVIGATION OWN

3. **BAW861** OWN NAVIGATION **BAW861** DIRECT_TO RAPET none

Contents

1. Overview of Assistant Based Speech Recognition in ATM
2. The view of an ANSP:
Results from Trials with
Nuance Recognition Engine



ANS CR - Operational View on ASR

Today ATM technology is „rich“ ..

- Electronic and interactive environment (not only radar screen)
- ATC support tools (AMAN/DMAN, A-CDM, datalink,..)
- Safety nets (STCA, APW, MTCD..)

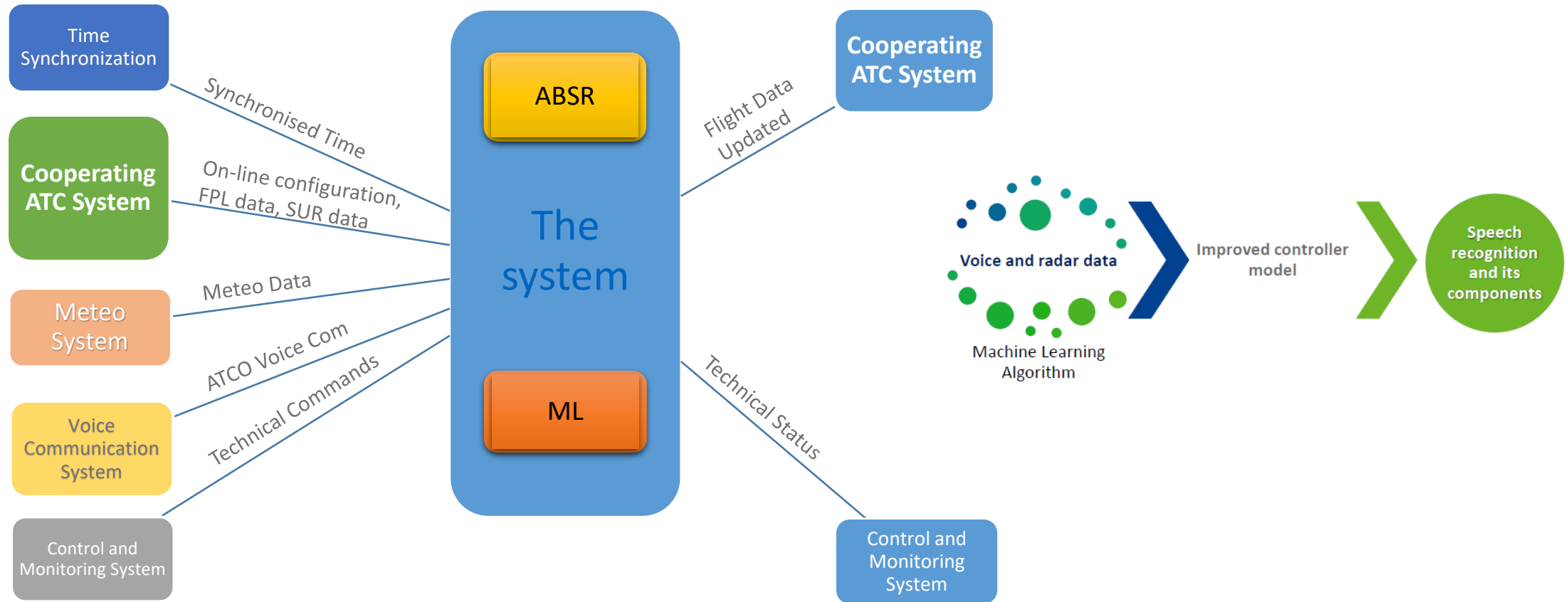
..and may provide better service.

End-user inputs are required/needed: correct, complete, on time,..

End-user **workload** is closely monitored.

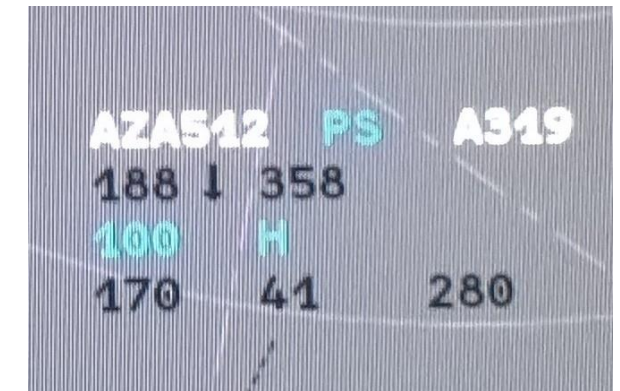
- Automation and integration of data
- ASR to be introduced in the operational room

Voice Recognition in ATM Environment



PJ16-04 – Prague Approach

- Two validation scenarios (24 and 30 MVT/h);
- 1 ATCo in executive position (TMA);
- Automatic highlight of controlled aircraft (*Labels*);
- Automatic input of all clearances and commands (speed, lateral and vertical movements, destinations and handovers);
- Input of data without an individual confirmation;
- Comparison of ATCo's workload and productivity (voice vs. PC *mouse*);
- Safety analysis for ASR (inserting data to the ATM system).



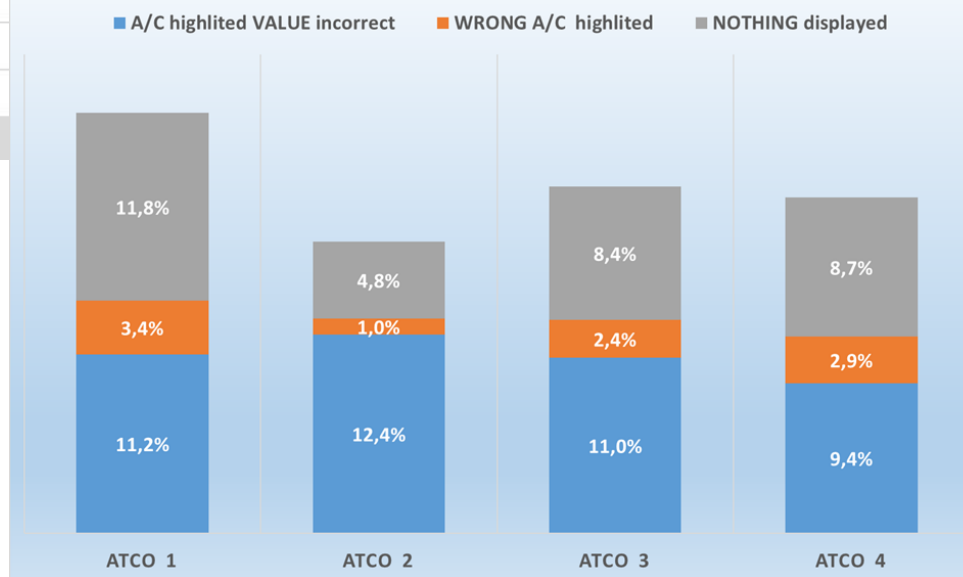
PJ16-04 – Prague Approach - RESULTS

	Total	RecognRate	ErrorRate
Altitude Cmds	581	82.10%	2.41%
Speed Cmds	164	76.22%	9.76%
Heading Cmds	324	63.89%	10.19%
Direct to waypoint	103	58.25%	15.53%
Cleared ILS	91	28.57%	0.00%
Handover Cmds	371	77.09%	1.08%
NO_CONCEPT	6	0.00%	0.00%
Initial Contact	218	86.70%	0.00%
QNH, Expect ILS	131	26.72%	0.00%
Sums	1989	70.64%	

Recognition Rate per ATCO (HMI)

	Scenario M	Scenario H	Both Scenarios
ATCO 1	76,3%	70,8%	73,7%
ATCO 2	78,5%	84,3%	81,7%
ATCO 3	80,1%	76,5%	78,3%
ATCO 4	78,1%	79,7%	79,0%
			78,3%

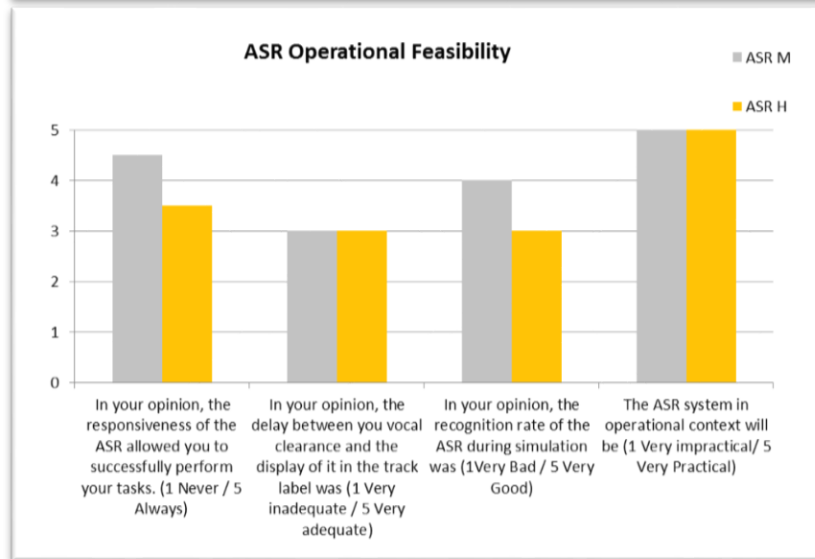
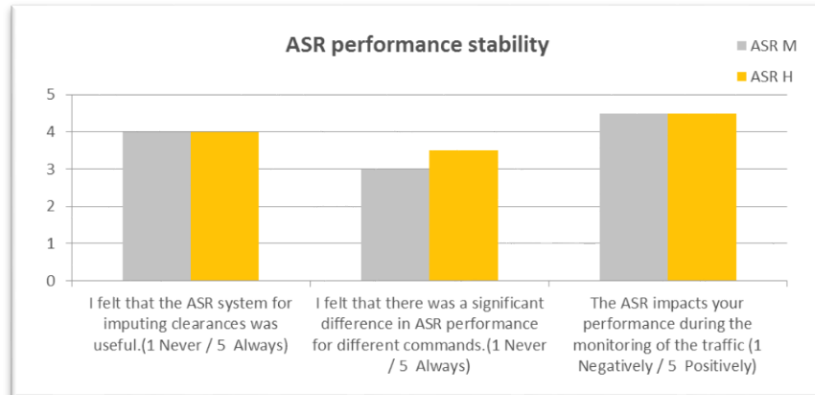
Error Rate per ATCO and Error type (HMI)



ond Run

PJ16-04 – Prague Approach - RESULTS

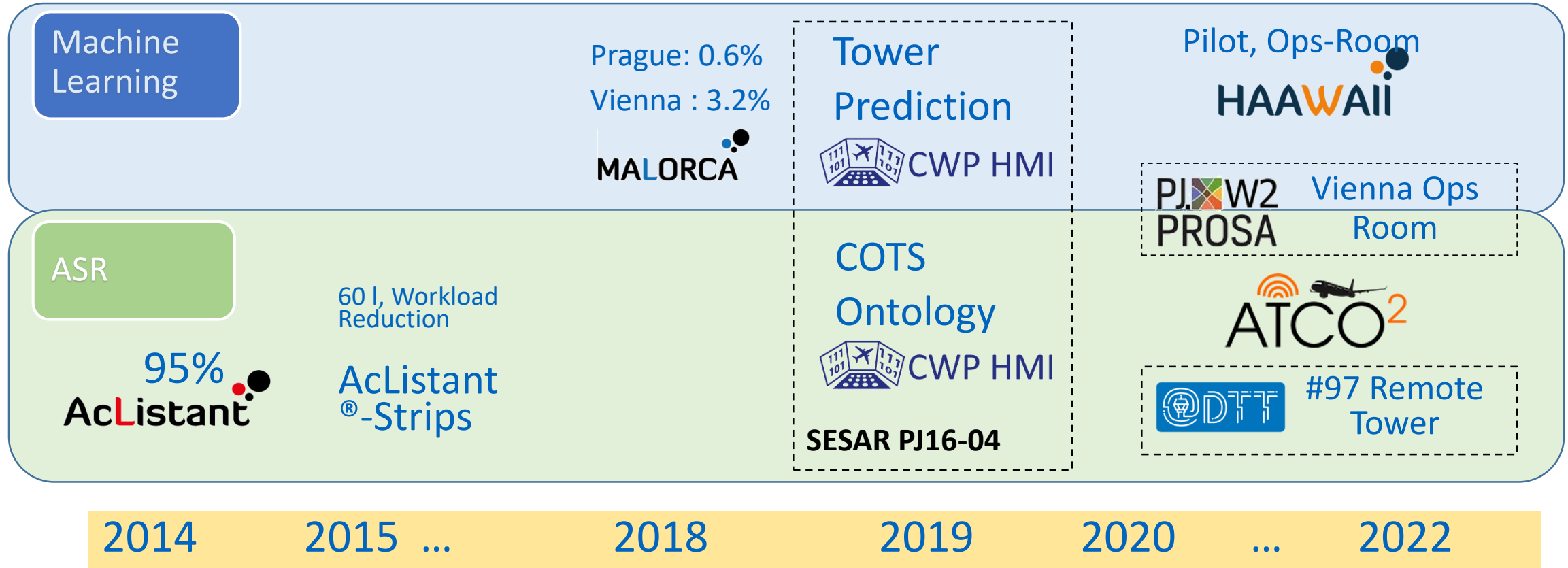
R/R Recognition Rate
E/R Error Rate
H Heavy



Interesting facts

- All 4 ATCos rated ASR as a **very practical system**;
- Although R/R was lower and E/R higher, the impact on ATCo's performance was assessed **positively** (4.5 out of 5);
- Automatic insertion of values into the Label, without explicit confirmation, has **acquired well**;
- Acoustic model for local English pronunciation is **necessary**;
- R/R was always higher in the 2nd run (*even for ATCos with Scenario H*) = ATCos intuitively **improved the pronunciation** to get support from ASR.

Speech Recognition and Machine Learning Roadmap



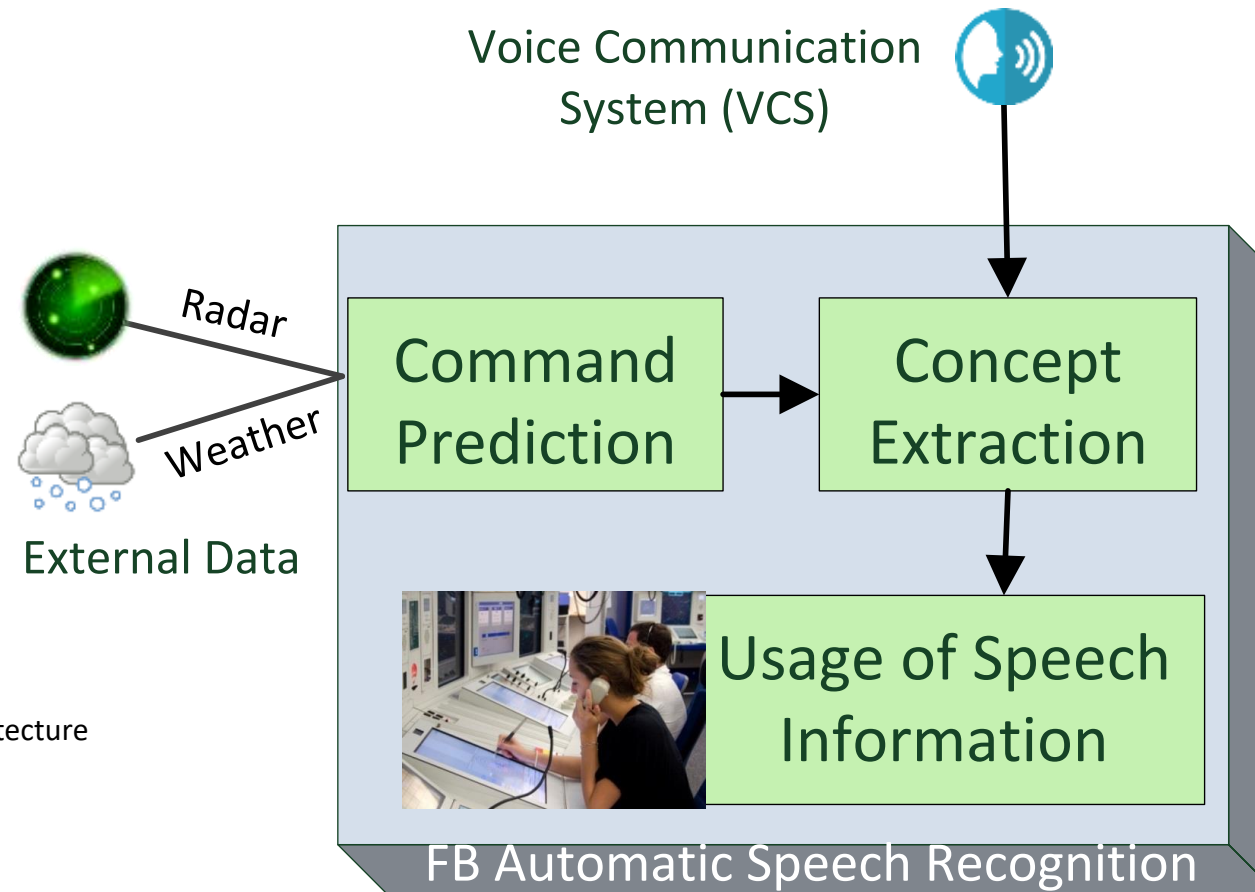


Thank you very much for your attention

Hartmut Helmke, Project Leader, DLR
Aneta Cerna, Project leader, ANS of the Czech Republic
Petr Hlousek, Project leader, ANS of the Czech Republic



Integration of **Speech Recognition** Into EATMA



EATMA = European Air Traffic Modelling Architecture
ASR = Assistant Based Speech Recognition
ABSR = Automatic Speech Recognition