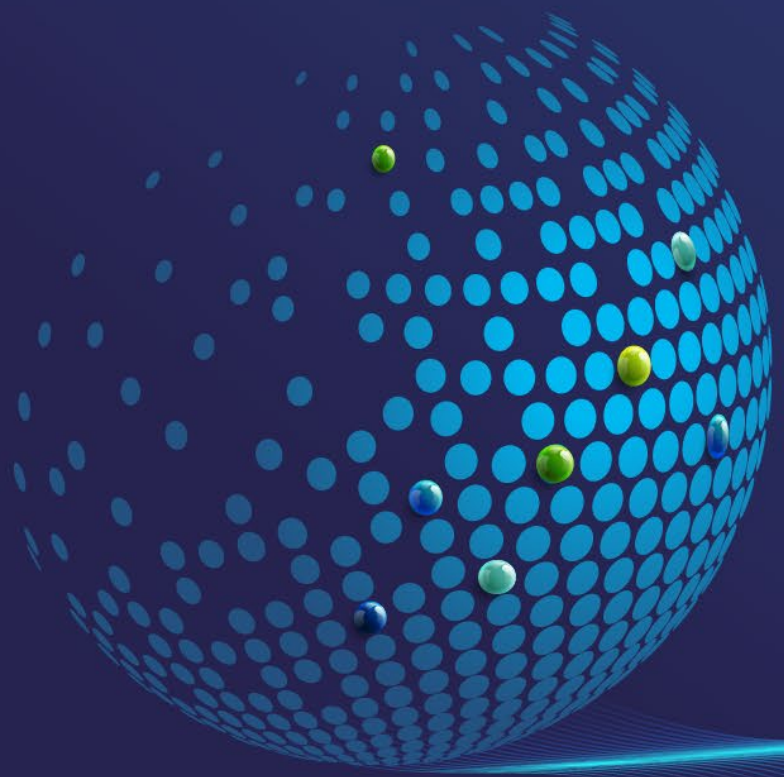


BEAM ME UP SCOTTY: INTELLIGENT HUMAN-MACHINE INTERFACES FOR THE DIGITAL AGE

9 March 11:00 – 12:00

SESAR 2020 SHOWCASE

#SESARShowcase

#EuropeForAviation

Automatic Speech Understanding in ATM Tower and En Route Environments

SESAR Validations of a natural complement to ATCOs HMI

Walter Rinaldi – Leonardo S.p.A.

Speech Recognition and Understanding in ATM

Differences from consumer services, specifics

- A few issues to overcome:

- Alexa, Siri, Google assistant not possible in ATM: (security, data protection, ATM data to third parties, external networks…)

- Delay and deterministic latency are very important (real time processing, predictable delays)

- Acquired ATM data volumes not large. Going to grow and improve, but not with as for Siri, Alexa… (GDPR, security, limited user base)

- Error resolution: ATCOs to find errors and resolve them. Crucial Aspect in ATM with safety implications

- The good news:

- ATM has a very limited and structured phraseology

- Using English only can help in the early stages of adoption

- ASRU does not require vast processing power

Uses of ASRU in ATM

- Can it work in ATM?
  - Yes, as several research projects proved in a number of instances (DLR team stands out)
- What can it do?

For ATCOs

- expedite/automate ATM tasks and reduce workload

- specific support for Datalink communications

- replace typing and mouse clicks

- read-back

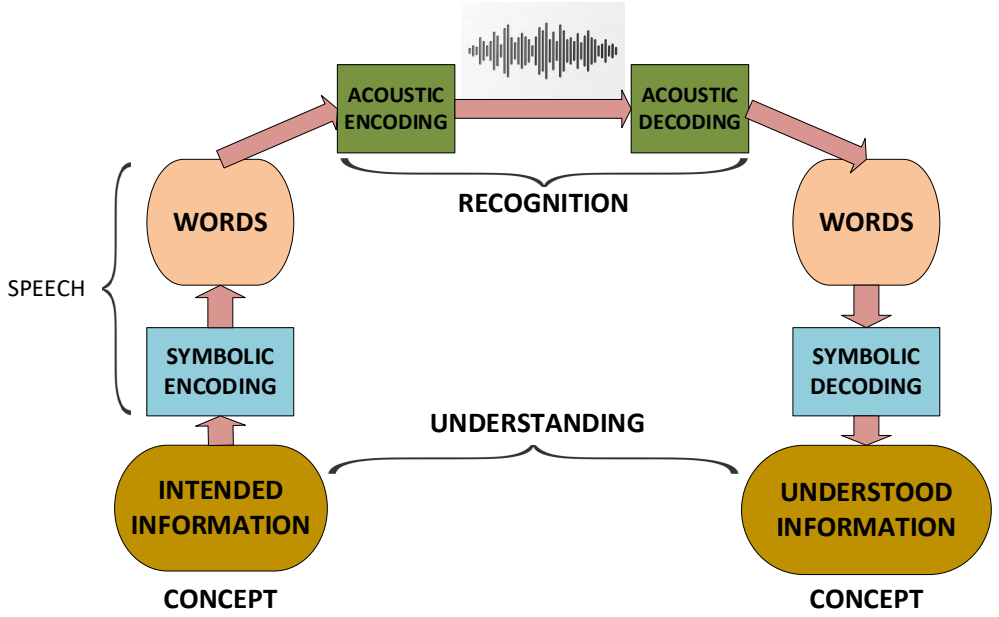
- help for training and simulation

For Pilots

- replace keying in/typing

- help with training and simulation

UNDERSTANDING MEANS DOING MORE THAN RECOGNIZING

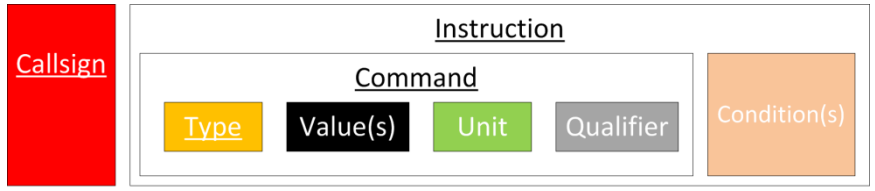


ASRU in ATM voice communications

- "hello lufthansa nine eight five descend flight level mmmh one hundred"

  DLH985 DESCEND FL 100

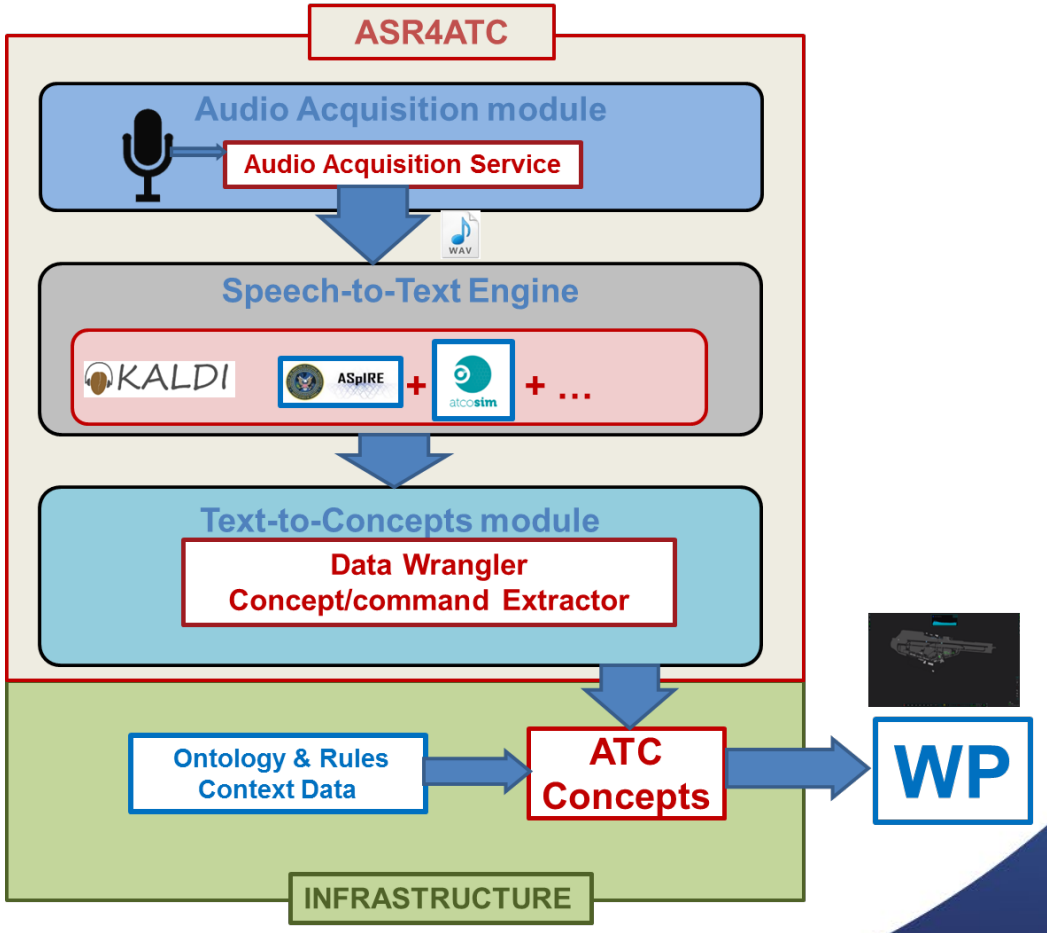
  words in green must be discarded
  (ASRU has still to recognize non relevant information, in order to discard it)



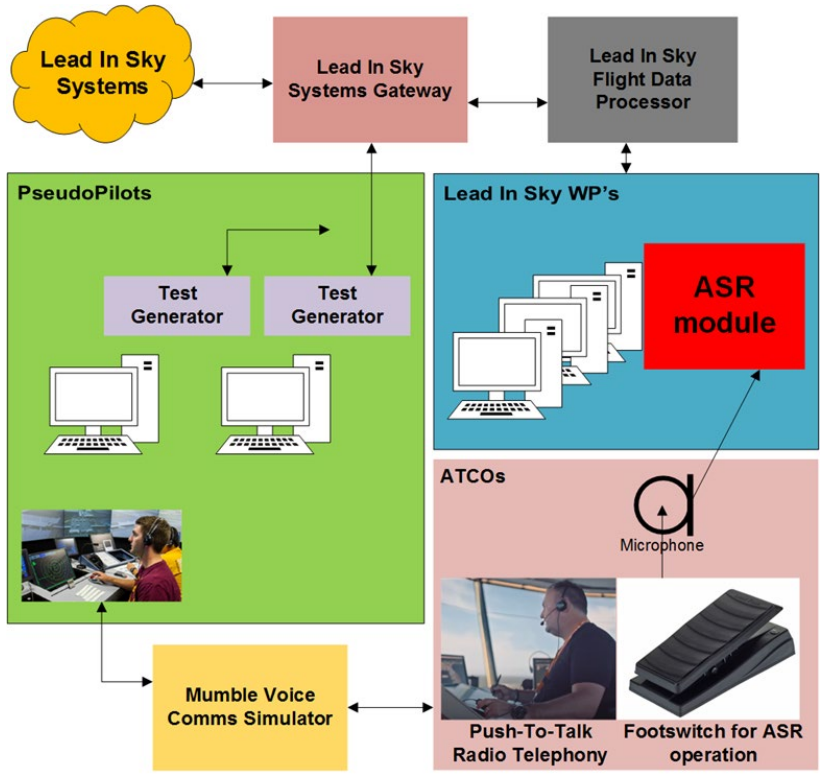
- Ontology (Agreed between European Air Traffic Management Stakeholders) and set of transcription rules: important to maximize data and design reuse
- Simplest use case, yet most appreciated: Highlighting flights

Leonardo SESAR ASRU Implementation

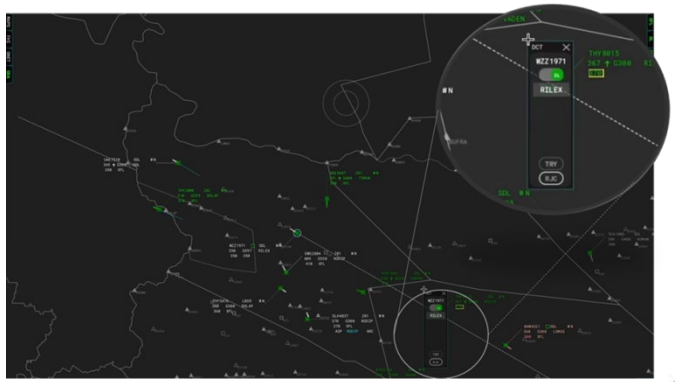
- Sesar PJ05 and 10 offered good entry points for ASRU exploitation
- Need for a quick solution, maximizing design reuse
  - audio acquired from microphone via a Java widget
  - trained phonetic models extract syllables/words
  - recognized word are transcribed (speech to text)
  - mapping via ontology based on ATM phraseology
  - concept extraction (text to concept) via ML
    - Based on a Fixed Vocabulary
    - Data manually transcribed and annotated
    - Available English language corpora
  - pre-trained model accounts for
    - -Language accents
    - -Phraseology deviations
    - Environmental limitations and noise
- CONTEXT-BASED DATA CRUCIAL TO ACHIEVING GOOD RESULTS
- DEVELOPED IN ONE YEAR FROM SCRATCH



Integrating ASRU with Leonardo Lead In Sky



Ground



Air

- Footswitch activates ASR
- ATM Ground/Tower/Air environments
- Basic command range
- Pared down ontology
- Strict observance of ICAO phraseology
- Callsign list of active/live flights used as context-based data
- Commands implemented on Lead In Sky Ground and Air WPs

Leonardo ASRU in SESAR Simulations

PERFORMANCE RATES:

- AIR: Callsign Recognition Rate ~ 80% Command Type Recognition Rate ~ 91%
- GND: Callsign Recognition Rate ~ 90% Command Type Recognition Rate ~ 76%

low error rates (no result is better than an incorrect result)

ATCOs ASSESSMENTS IN TERMS OF

- Workload
- Situational Awareness
- Potential for Human Error
- Trust
- Usabilty
- Technology Acceptance
- Safety

Resulting from questionnaires were generally positive (graphs omitted)

And In The End...