INDUSTRIAL RESEARCH

SESAR SOLUTION 97.2 CONTEXTUAL NOTE TRL4

Deliverable ID:	D3.1.095
Dissemination Level:	PU
Project Acronym:	DTT
Grant:	874470
Call:	H2020-SESAR-2019-1
Topic:	SESAR-IR-VLD-WAVE2-05-2019
Consortium Coordinator:	AT-One consortium
Edition date:	23 December 2022
Edition:	[00.01.00]
Template Edition:	02.00.04

EUROPEAN PARTNERSHIP

Co-funded by the European Union

Authoring & Approval

Authors of the document

Beneficiary	Date
ENAV	15 November 2022

Reviewers internal to the project

Beneficiary	Date
DLR (AT-ONE)	1 December 2022
NLR (AT-ONE)	1 December 2022
ANS CR (B4)	9 December 2022 (S)
ON (B4)	9 December 2022 (S)
ACG/ COOPANS	9 December 2022 (S)
CCL/ COOPANS	9 December 2022 (S)
LFV/ COOPANS	9 December 2022 (S)
ENAIRE	15 December 2022
HC (FSP)	9 December 2022 (S)
INDRA NO	6 December 2022
INDRA	5 December 2022
LDO	1 December 2022
INTEGRA	20 December 2022 (S)
DBLUE	15 December 2022
UNIBO	9 December 2022 (S)

Reviewers external to the project

Beneficiary	Date
EUROCONTROL	9 December 2022 (S)
LFV	9 December 2022 (S)

Approved for submission to the S3JU By - Representatives of all beneficiaries involved in the project

Beneficiary	Date

Rejected By - Representatives of beneficiaries involved in the project

Beneficiary	Date

Document History

Edition	Date	Status	Beneficiary	Justification
00.00.01	15 November 2022	First draft	ENAV	First version for partners revision
00.01.0o	23 December 2022	Version for submission	ENAV	Comments included

DTT

DIGITAL TECHNOLOGIES FOR TOWER

This Contextual Note is part of a project that has received funding from the SESAR Joint Undertaking under grant agreement 874470 under European Union's Horizon 2020 research and innovation programme.

Abstract

This TRL4 contextual note illustrates all relevant aspects of SESAR Solution 97.2 in support of Solution TRL-4 Maturity Gate.

Solution 97.2 "Automatic speech recognition (ASR) supported by AI and Machine Learning", investigates the use of new human-system interaction in ATC tower environment, enabled by automatic recognition and translation of spoken language into the ATC system, with the aim to reduce ATCO workload and improve safety.

Table of Contents

List of Tables

List of Figures

Non è stata trovata alcuna voce dell'indice delle figure.

1 Purpose

This contextual note introduces SESAR Solution 97.2 with a summary of the results stemming from R&D activities conducted. It provides to any interested reader (external and internal to the SESAR programme) an introduction to the SESAR Solution in terms of scope, main technical and performance benefits, relevant system impacts as well as additional activities to be conducted during the next validation phase. This contextual note complements the technical data pack comprising the SESAR deliverables required for TRL4 Maturity Gate and further validation phases[1].

1 "The opinions expressed herein reflect the authors' view only. Under no circumstances shall the SESAR3 Joint Undertaking be responsible for any use that may be made of the information contained herein."

[2] Applicable Integrated Roadmap dataset is DS23.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

2 Improvements in Air Traffic Management (ATM)

The solution "ASR at the TWR CWP supported by AI and Machine Learning" addresses the development of new human machine interface (HMI) interaction modes at the Tower Controller Working Position thanks to the automatic speech recognition and translation of spoken language into the system input, to automatically supporting some ATCO tasks, which are either performed manually or not performed at all in currently available systems/CWPs, with the aim to reduce ATCO workload and improve safety in several airport sub-operating Environments.

An ASR engine is able to record verbal utterances from the controller-pilot communications, recognise a list of word (transcription), which is the basis for the extraction of concepts, i.e. extraction of ATC commands following the defined ontology (annotation). The extracted controller commands constitute a hypothesis, which needs to be checked against contextual information, eventual manual inputs, and a set of possible commands given by the "Command Hypotheses Predictor" function, basing on a machine learned command prediction model on historical surveillance and speech data. The finally checked controller commands, i.e. the most reasonable hypotheses due to the Assistant Based Speech Recognition (ABSR) functionality chain, can then be used for further ASR applications. Identified use cases in the scope of the solution are:

- The recognition of verbal controller-pilot communication and highlighting of recognized callsign;
- The recognition of verbal controller-pilot communication and display of full recognized utterance/command in HMI;
- The manual manipulation of an ASR output in case the controller spots any misinterpretations.
- The automatic acceptance of ASR output within an (adjustable) time frame (e.g. 10 seconds).

These improvements have been validated in simulated Small/Other and Medium size airports, while they may be applicable in all airport sub-operating Environments.

3 Operational Improvement Steps (OIs) & Enablers

This Solution is linked to a new OI, the POI-0040-SDM named **"Automatic Speech Recognition with AI/ML at the TWR CWP"**, composed of the following technologies:

- Introduction of new automated functions for Automatic Speech Recognition using AI and Machine Learning Techniques at the Aerodrome CWP/HMI for improving the controller workload.

Technologies associated with Solution 97.2 have been defined by Solution team and submitted for endorsement in the Master Plan[2].

OI Step code	Title	Description
POI-0040-SDM	Automatic Speech Recognition with AI/ML at the TWR CWP	ATCOs will be supported by introducing innovative human machine interaction such as Automatic Speech Recognition that can be enhanced by the use of Machine Learning. The goal is to automatically support certain tasks of the ATCO, which are not done or done manually today in today's systems/ CWPs.

Table 1 Solution relevant OIs

Enabler Code	Title	Description	Background	Coverage
AERODROME-ATC-106	Automatic Speech Recognition supported by AI and ML algorithms for aerodrome tower operations	Introduction of new automated functions for Automatic Speech Recognition using AI and Machine Learning Techniques at the Aerodrome CWP/HMI for improving the controller workload.	16.04	Full

Table 2 Solution Enablers

[2] Applicable Integrated Roadmap dataset is DS23.

4 Background and validation process

Previous SESAR work provides the baseline for ASR related research. The following work and related recommendations for improvement has been considered:

- MALORCA project focused on the Automatic Speech Recognition, aiming to significantly reduce controller's workload and increase ATM efficiency. Among the project achievements, command error rates comprised between 2% and 5%, resulting in command recognition rates between 90% and 95%. These recognition rates were achieved thanks to the so-called assistant-based speech recognition, dynamically generating context information to increase the recognition rate.

- In SESAR 2020 Wave 1, the Industrial Research project PJ.16-04 dealt with new methods of controller interaction with the Human Machine Interface (HMI) at the Controller Working Position (CWP). The solution developed guidance and assessment methods regarding HMI, investigated new HMI needs and interaction modes in relation with SESAR solutions (including new user interface technologies such as speech recognition, multi-touch, and gaze detection). Results of this solution have been capitalised as a valuable baseline for Solution 97.2.

In Wave 2, the research on Automatic Speech Recognition applications for ATC Tower has been carried on in the frame of Solution 97.2, through a set of three technical validations:

- One Real Time Simulation addressing Speech Recognition in a multiple remote tower environment performed by INDRA and HUNGAROCONTROL in Asker.
- One Real Time Simulation addressing Speech Recognition at Braunschweig simulating a multiple remote airport controller working position adapted from existing airports, led by DLR.
- One Real Time Simulation addressing Speech Recognition at Rome simulating Sofia airport, led by LEONARDO.

5 Results and performance achievements

SESAR2020's Solution 97.2 has extensively addressed the Automatic speech recognition and understanding for air traffic control (ATC) communication in simulated tower and ground environments, analysing the recognition rates and human performance of air traffic controllers (ATCos). Three validation exercises with 22 ATCos from four different European air navigation service providers were conducted in Germany, Norway, and Italy. The validated artificial intelligence-based prototypes of Assistant Based Speech Recognition systems (ABSR) supported ATCos in fulfilling tasks in a ground and tower environment as well as multiple remote tower environment, respectively. Thus, in any relevant ATC display, (1) recognized callsigns of ATCo utterances have been highlighted, (2) fully recognized commands were shown, and (3) the ATCo was able to manually manipulate the ABSR output if needed or the output was automatically accepted by the ATC system otherwise.
The three validations which took place in Solution 97.2 proved the technical feasibility of the ASR technology to capture Aerodrome ATC instructions and clearances transmitted by radio to flight crews and to use them to automate ATC system inputs.

The main findings from the overall validation exercises can be summarized as follows:

- In general, ATCOs saw the potential in applying speech recognition in a TWR environment and were able to perform their ATC tasks (even given the CWP prototypic systems) when working with ASR support. The outcomes indicated that ASR has no negative impact in terms of workload and situation awareness and therefore do not appear to reduce safety levels, while the positive results for system usability, job satisfaction and some workload measurements show the potential of ABSR in a (multiple remote) tower environment and foster to go further in maturity level.
- Encouraging feedback from ATCOs regarding acceptance and trust in the system indicate that the level of ASR technical performance was acceptable and consistent with human capabilities.
- Different results of recognition rates were collected at the three validations: a callsign recognition rate of 81-98%, a command recognition rate of 65-91%, and a slight reduction in ATCo workload on a low workload level.
- The quantitative and qualitative feedback of ATCOs were good and motivating to go beyond TRL4 and would have been even better if the full potential of ABSR accuracy have been offered to them.
- Positive reactions from ATCOs in terms of usability suggest a high quality of user experience when interacting with ASR and its related functions. Nevertheless, some degree of training would be required for ATCOs to better understand "behaviours" of ASR and also to learn how to proactively adapt their speech to the tool.
- The "hook" function (the highlight of the label of aircraft addressed, as soon as it is pronounced by controller) was found to bring benefit to tower controllers' situation awareness.
- Safety aspects were addressed across all runs and no specific safety issues were identified during the validation exercise.
  The data shows that ATCos speak differently, i.e., closer to phraseology if being supported by ABSR (i.e., solution runs have higher command recognition rates that baseline runs; in the latter, the speech was analysed as well, but the output was not shown to the ATCo). On one hand, this might be, because they get better support if recognition rates are higher, on the other hand, it might be due to the pure awareness of working with speech recognition in the

background. If ATCos are sticking closer to ICAO phraseology just by pure presence of an ABSR system, that could already be a safety feature.

6 Recommendations and additional activities

At the end of the Technical Validations on ASR, the following recommendations are made to support the future steps of validation of this Solution.

The European-wide agreed ontology for annotation of ATC utterances developed in the ASR projects built a base for interoperability of systems. Nevertheless, the vocabulary could be extended, enriching the type and number of ATC commands, or considering the benefit to automate an input (such as assuming traffic or activation/de-activation of stop bars). Making callsign range wider, including military, GA, more formats and airline operators has been recommended.

The ASR for pilot side could be implemented to highlight callsign when a pilot is calling in or, instead of automatically updating the EFS, the system could function in a more “preventative” manner and check whether the pilot provided the correct readback and notify the ATCO in case of a mismatch. Such an “error prevention” functionality could have a positive impact on safety and overall end-user acceptance.

The use of Machine Learning could be improved to provide better interpretation of ASR.

Appropriate training of the ATCOs to the new system functionalities should be provided in order to achieve better understanding and to build up trust into the new functionalities.

Further possible recommendation for future phases have been identified as possible uses of ASR:

- To be further developed as a potential help to on job training.
- To be further developed as potential help for incident analysis (e.g. offline tool to analyse transcriptions).
- Allowing controllers to search a/c also by their type would significantly improve the ‘Hook’ function effectiveness.
- To foresee an interaction between ASR and the eFlight Strips, to allow ATCOs to activate, via ASR, safety barriers such as stop bars.
- To foresee a further and tighter integration with A-SMGCS functionalities. For example, it would be useful if ASR could recognise and display the taxi route assigned to an a/c by a controller or if it could display a runway as ‘occupied’ when recognising that a vehicle using that runway is in contact with the tower. Also, ASR could highlight a closed taxiway on the WP HMI.
- To increase callsign recognition capability including military and general aviation callsigns;
- To implement voice commands specific for the interaction with system software menus (e.g. what-if; lists; interaction with windows or graphic objects; TCT, etc.)
- To integrate updated artificial intelligence (AI) and machine learning algorithms for the intelligent speech recognition and data cross check, evolving what done in previous phases.

Display design could be improved to increase situational awareness, in terms of background colours of the ASR pop-up window, visibility of strip and ‘Hooked’ a/c highlight. Introducing an ‘ASR pop-up window’ would display logs and transcripts, always in the same place, similar to a chat window, in order for ATCOs to inspect ASR operation when and if necessary.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

The amount of training data must be further improved given representative samples, i.e., the portion of female voices in ATC communication data is much less than of male (also leading to worse results of recognition rates). Furthermore, a big amount of data must be recorded from operations rooms (not from labs), because ATCOs speak differently in simulations.

7 Actors impacted by the SESAR Solution

Tower Clearance Delivery Controller, Tower Ground Controller, Tower Runway Controller.

8 Impact on Aircraft System

This solution investigates technologies for tower and has no impact on Aircraft Systems.

9 Impact on Ground Systems

The Technological Solution will consist of additional systems to complement the CWP HMI currently used.

The Tower controller working position will be equipped with the Assistant Based Speech Recognition (ABSR) prototype, which enables the system to recognize the callsigns and commands given by the air traffic controller for callsign highlighting and automatic input of commands into system input. The air traffic controller utters ATC instructions after pushing the push-to-talk-button on a headset (or a pedal) until releasing the button again. The audio stream is continuously analysed by an ASR engine, i.e., the speech-to-text functionality already delivers word sequences before the air traffic controller finishes the audio transmission and subsequently the word sequences are turned into ATC concepts, so that the system automatically issues commands given by ATCOs.

10 Regulatory Framework Considerations

SESAR2020 W1 PJ.16-04 developed an ontology for the transcription of controller commands (En-Route, approach, tower) as well as for the hypotheses input and output transcription standard, that has been shared and agreed among some of the major European ANSPs, ATM system providers and research institution and that can be treasured as a good basis for future proposals on standardisation of content and format, i.e.: speech-to-text with a number of word sequence hypotheses, text-to-concept based on the ontology for ATC utterances and preparations in order to feed succeeding applications such as runway error detection, formats such as JSON for content transmission, and many aspects more to enable comparability and interoperability.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

11 Standardization Framework Considerations

Standard could be identified for the voice sampling rate (kHz).

The ontology for ATC commands has been further enhanced. The need for standardising the phraseology for the speech recognition within the TWR domain has been assessed as part of the solution activities. However, it has been agreed that there is not a strong need to convert the ontology into a standard, since sticking the phraseology to a mandatory standard would difficult the deployment of the concept throughout the different dependencies (for instance, there would be problems for using local languages).

As part of PJ.05-W2-97.2 activities, a communication of the findings and results of the Solution to EUROCAE Technical Advisory Committee has been carried out. As a result of this coordination, some guidance has been provided with regards to the standardisation needs for ASR:

- There is not an existing standard for Voice Recognition in the ATM environment. An assessment of this need should be performed in further stages of the developments.
- Outside the ATM environment, there is an existing standard that is relevant for the solution: ISO/IEC 30122-2:2017, which provides the technical criterions and test methods of voice commands and its speech recognition engine. It is recommended that this standard is taken into account when developing the ASR functionality for ATM.
- Proposals for standardisation of the content and the format for input and output of assistant based speech recognition systems should be identified, i.e., speech-to-text with a number of word sequence hypotheses, text-to-concept based on the ontology for ATC utterances and preparations in order to feed succeeding applications such as runway error detection, formats such as JSON for content transmission, and many aspects more to enable comparability and interoperability.
- The usage of commercial-off-the-shelf products is not feasible for the ATM environment. Therefore, dedicated products developed for this environment would match better the expectations and requirements for deploying the concept in ATM.

12 Solution Data pack

As from agreement with SJU, *one* unique Data Pack is provided for the *two* PJ.05-W2-WP3 Solutions 97.1 and 97.2, with dedicated paragraphs for each.

The D3.1 Solution PJ.05-W2-97.1 and 97.2 TRL4 Data Pack includes:

Solution Data Pack	
Systems Requirements consolidation	**Technical Specification/Interface Requirements Specification (TS/IRS) TRL4 - Final version** D.3.1.022 Edition date: 28/09/2022 Edition: 00.02.04
	TS/IRS Part II (SAR) D.3.1.022 Edition date: 03/10/2022 Edition: 00.00.04
	TS/IRS Part IV (HPAR) D.3.1.022 Edition date: 30/09/2022 Edition: 00.01.00
	PAR *(Additional document for technological solution, non PMP)* Edition date: 12/10/2022 Edition: 00.01.00
Cost Benefit Analysis tailored for the specific Technological Solution (CBAT) TRL4	**Cost Benefit Analysis (CBAT) Final - TRL4** D3.1.071 Edition date: 26/10/2022 Edition: 10.00.01
Technical Validation Report (TVALR) TRL4	**Technical Validation Report (TVALR) Final version – TRL4** D.3.1.051 Edition date: 15/11/2022 Edition: 00.13.00

Initial Technical Validation Plan (TVALP) defining the validation roadmap for TRL 6	**Initial Technical Validation Plan (TVALP for TRL6)** D.3.1.080 Edition date: 7/10/2022 Edition: 00.01.00

Table 3 Solution Data Package

Air Navigation Services of the Czech Republic

austro CONTROL

CRIDA I+D+i - ATM

CROATIA CONTROL

deepblue

DLR

ENAIRE

enav

HungaroControl

indra

integra

LEONARDO

nlr

ORO NAVIGACIJA

EUROPEAN PARTNERSHIP

Co-funded by the European Union