D4.1.120-PJ.10-W2-96 ASR TRL6 CONTEXTUAL NOTE

Deliverable ID:	D4.1.120
Dissemination Level:	PU
Project Acronym:	PROSA
Grant:	874464
Call:	H2020-SESAR-2019-1
Topic:	Separation Management and Controller Tools
Consortium Coordinator:	DFS
Edition date:	19 June 2023
Edition:	00.04.00
Template Edition:	02.00.04

EUROPEAN PARTNERSHIP

Co-funded by the European Union

Authoring & Approval

Authors of the document

Beneficiary	Date
LEONARDO	08/03/2023

Reviewers internal to the project

Beneficiary	Date
ANS CR	27/03/2023
CROCONTROL	27/03/2023
DLR (AT-One)	30/03/2023
ENAIRE/CRIDA	30/03/2023
INDRA	27/03/2023
INTEGRA	27/03/2023
NLR	27/03/2023
NATMIG/SINTEF	27/03/2023
AUSTROCONTROL	27/03/2023

Reviewers external to the project

Beneficiary	Date

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

Beneficiary	Date
ANS CR	Silent Approval
CROCONTROL	Silent Approval
DLR (AT-One)	31/03/2023
ENAIRE/CRIDA	30/03/2023
INDRA	Silent Approval
INTEGRA	Silent Approval
LEONARDO	27/03/2023
NLR	Silent Approval
NATMIG/SINTEF	Silent Approval
AUSTROCONTROL	Silent Approval

EUROPEAN PARTNERSHIP

Co-funded by the European Union

Rejected By - Representatives of beneficiaries involved in the project

Beneficiary	Date

Document History

Edition	Date	Status	Beneficiary	Justification
00.00.01	08/03/2023	Draft	LDO	Draft version for partners's review
00.00.02	27/03/2023	Draft	LDO	Reviewed version and reference to Malorca project.
00.00.03	30/03/2023	Draft	ENAIRE/CRIDA	Review and contribution to section 9 and results
00.00.04	30/03/2023	Draft	DLR	Review and comments and changes to all pages
00.01.00	31/03/2023	Final	LDO	Final version for submission
00.02.00	20/04/2023	Final	LDO	Final version for submission addressing SJU comments
00.03.00	05/06/2023	Final	LDO	Final version addressing the outcomes of the Maturity Gate.
00.04.00	19/06/2023	Final	LDO	Final version addressing SJU comments to previous version.

Copyright Statement

EUROPEAN PARTNERSHIP

Co-funded by the European Union

PJ.10-W2-PROSA

SEPARATION MANAGEMENT AND CONTROLLER TOOLS

Solution 96 ASR is part of a project that has received funding from the SESAR3 Joint Undertaking under grant agreement No 874464 under European Union's Horizon 2020 research and innovation programme.

Abstract

Automatic Speech Recognition or ASR, as it is known in short, is the technology that allows users to utilize their voices to speak with a computer interface in a way that allows converting human speech into texts, extract the information within, and use this knowledge in different applications. The ASR enabled applications can provide a solution to significantly reduce air traffic control operators' (ATCOs') workloads and increase ATM efficiency.

Considering the continuous growth of air traffic, ASR is a technology that can help the controller to carry out their task in a more efficient way.

SESAR Solution PJ.10-W2-96 ASR deals with new methods of controller interaction with the Human Machine Interface (HMI), applying mature technologies from other domains to ATM. This will increase controller productivity, reduce workload, reduce stress level and enable the use of SESAR advanced tools, safely facilitating performance-based operations.

Table of Contents

Abstract 4
1 Purpose 6
2 Improvements in Air Traffic Management (ATM) 7
3 Operational Improvement Steps (OIs) & Enablers 8
4 Background and validation process 10
5 Results and performance achievements 11
6 Recommendations and Additional activities 13
7 Actors impacted by the SESAR Solution 14
8 Impact on Aircraft System 15
9 Impact on Ground Systems 16
10 Regulatory Framework Considerations 17
11 Standardization Framework Considerations 18
12 Solution Data pack 19

List of Tables

Table 1 Operational Improvement steps and Enablers for Solution 96 ASR 8

List of Figures

No table of figures entries found.

1 Purpose

This contextual note provides to any interested reader (external and internal to the SESAR programme) an introduction to the SESAR Solution 96 ASR in terms of scope, main operational and performance benefits, relevant system impacts.

As the Solution is at TRL6 level it contains as well additional activities to be conducted during the industrialization phase or as part of deployment.

This contextual note complements the technical data pack comprising the SESAR JU deliverables (for TRL6, they are proposed to support industrialization/deployment).

EUROPEAN PARTNERSHIP

Co-funded by the European Union

2 Improvements in Air Traffic Management (ATM)

Solution PJ.10-W2-96 ASR supported by AI and Machine Learning addresses the development of new human machine interface (HMI) interaction modes at the Controller Working Position aiming at introducing Automatic Speech Recognition in the ATM domain, as a means to improve Air Traffic Control Operators's (ATCO's) situational awareness, reduce workload and increase productivity. Several commands can be recognised from the controller-pilot spoken dialog, and they can be directly presented (and input) on the Controller HMI instead of manually inputting them into the system. Highlighting of targets upon callsign pronunciation from the ATCO or pilot is also considered, as well as the use of functions in the CWP such as visualisation of sectorisation changes and 3D navigation. Artificial intelligence (AI) and machine learning algorithms are required for speech recognition engine training.

An ASR engine is able to process 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, can then be used for further ASR applications. Identified use cases in the scope of the solution are:

- Highlight of callsigns on the CWP from pilot utterances;
- Highlight of callsigns on the CWP from controllers' utterances;
- Annotation of controllers' commands;
- Pre-filling of commands in the CWP;
- Voice commands for highlighting an upcoming sectorization change in the CWP;
- Voice commands for highlighting the fights that will be affected by an upcoming sectorization change in the CWP;
- Voice commands for navigating the 3D visualization of the air space in the CWP;
- Prefilling of Datalink commands.

These improvements have been validated in a TMA and En-Route with low to medium complexity and high workload operating environment, while they may be applicable in all operating Environments.

Not all the concepts have reached TRL6, some are still TRL6 ongoing as decribed later with the need of splitting the original enabler into free new one.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

3 Operational Improvement Steps (OIs) & Enablers

The Solution is an enabling solution in the S2020 framework and covers the SESAR Roadmap OIs and Enablers described in the table below. The corresponding Change Request will be endorsed in (DS23).

OI Step	OI description	Open CR
POI-0055-SDM	Improving controller productivity by Automatic Speech Recognition (ASR) at the ER/APP CWP/HMI.	No CR opened in Wave 2, the POI-0055-SDM has been created in Wave 1.
EN code	**EN description**	**Open CR**
ER APP ATC 180	Controller productivity enhancements by Automatic Speech Recognition at the ER/APP CWP/HMI	No CR opened in Wave 2, the Enabler ER APP ATC 180 has been created in Wave 1
	Deletion of the Enabler ER APP ATC for the need of splitting it into three new enablers	CR 07534 opened for deleting the enabler ER APP ATC
ER APP ATC 227	Use of Automatic Speech Recognition to recognise commands from ATCO utterances	CR 07531 opened for creation of the new enabler
ER APP ATC 228	Use of Automatic Speech Recognition for local and configuration inputs at the ER/APP CWP/HMI	CR 07532 opened for creation of the new enabler
ER APP ATC 229	Use of Automatic Speech Recognition to highlight flights by recognition of callsigns	CR 07533 opened for creation of the new enabler

Table 1 Operational Improvement steps and Enablers for Solution 96 ASR

After the execution of the validation exercises and their outcomes, showing different maturity level achieved in the specific applications of the ASR, the Solution upon suggestion of the SJU has split the enabler ER APP ATC 180 in the following three ones:

Use of Automatic Speech Recognition to recognise commands from ATCO utterances (TRL6 achieved)

Application of ASR to process commands from ATCO utterances, such as cleared flight levels, direct routings or headings, to input them into the En-Route / Approach ATC ground system and, as a result, update the system flight plan. The trajectory might be updated as well, and CPDLC messages sent, if applicable. ASR is supported by Machine Learning and Artificial Intelligence.

Use of Automatic Speech Recognition for local and configuration inputs at the ER/APP CWP/HMI (TRL6-ongoing)

EUROPEAN PARTNERSHIP

Co-funded by the European Union

Application of ASR to manage functionality local to the CWP, essentially related to display and configuration, such as 3D navigation or highlight of sectors. ASR is supported by Machine Learning and Artificial Intelligence.

Use of Automatic Speech Recognition to highlight flights by recognition of callsigns (TRL6-ongoing)

Application of ASR to pilot and ATCO utterances to recognise the callsign of the aircraft calling and highlight the related flight at the ER/APP CWP. ASR is supported by Machine Learning and Artificial Intelligence.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

4 Background and validation process

PJ.10-W2-96 ASR starts taking into account the work performed by S2020 PJ16-04 Wave 1 project, as well as the MALORCA project, executed in the context of Exploratory Research. PJ.10-W2-Sol.96 ASR's starting maturity level is TRL4 and it targets to reach TRL6 maturity at the end of Wave 2 activities.

Automatic Speech Recognition matured up to TRL-4 in PJ.16-04 Wave 1 and enabled the definition of rules for transformation of a sequence of ATC words into ATC concepts (by means of the so called "ontology") and exploratory usage. PJ.16-04 concluded that speech recognition engines needed adaptation to achieve acceptable results.

PJ10-W2-96 ASR Solution also benefits from the experience in a number of other projects on Automatic Speech Recognition in ATC. The first assistant-based speech recognition system (ABSR) for an ATC approach area has been built in the project AcListant® with laboratory speech data of ATCos and validated for aircraft radar label maintenance task in the project AcListant®-Strips. However, the development of this first ABSR system based on expert knowledge and adapting it to different environments manually would have been costly. Thus, the idea of the following SESAR Exploratory Research project MALORCA for other ATC approach areas with ops room data from ATCos was to automatically learn the needed models, i.e., acoustic model, language model, command prediction model, etc. MALORCA project focused on the Automatic Speech Recognition, with the aim to significantly reduce controller's workload and increase ATM efficiency. The results demonstrated a command recognition rate between 90% and 95% (error rate between 2% and 5%) with assistant-based speech recognition (i.e. an AMAN dynamically generates context information to increase the recognition rate). Without context generation the recognition rate was only between 50% and 80%. Besides, the speech recognition was deemed to be the more reliable input sensor, at least in the simulation setup.

PJ.10-W2-96 ASR aims to improve the speech recognition engines adaptation and use real data in an operational environment. Starting from the consolidated ontology developed in PJ.16-04. PJ.10-W2-96 ASR tries to improve the recognition engine using context data.

Furthermore the Solution aims at integrating the speech recognition engine in an operational En-Route and TMA environment exploring how the use of the recognition engine can improve and facilitate the routinely tasks of ATCOs. The Solution has been evaluating the benefits of such a technology to reduce the ATCO's workload by highlighting the Callsigns of flights upon ATCOs and pilot utterances. Validations have also covered use cases such as Prefilling command masks in CWP (Controller Working Position) and also for datalink messages and investigating how ASR may be used to enable faster and more predictable navigation in 3D visualizations of the airspace sectorisation when using dynamic airspace configuration (DAC).

5 Results and performance achievements

In general Assistant Based Speech Recognition, based on Automatic Speech Recognition techniques, has established itself as a reliable support tool for many Air Traffic Management environments, and the validation exercises seem to bring encouraging results. Assistant Based Speech Recognition as first validated in AcListant® project combines a speech recognizer with an assistant system.

The ASR prototypes were successfully integrated with operational systems, with no negative impact on the other systems of the CWP and the performance of the previous system.

The command recognition rate was in line with the expectations for some exercises and a bit lower for other depending on callsign or command recognition, nevertheless the success criterion with differences in error rate smaller than 2.5% between different command types was not satisfied.

It was agreed during the Maturity Gate that this requirement is too stringent if applied to all commands. For example, it would be expected that the wrong recognition of a number would be more critical than the failure to recognize a DCT command due to the system not recognizing the waypoint the aircraft has been cleared direct to. Moreover, the benefits of speech recognition are likely to be different for types of commands (e.g. if entering the command in the HMI is easy, like in the case of a DCT, then the potential benefit of using ASR are lower).ATCOs stated that the workload did not change or was decreased using the ASR technology, they considered that the recognition rate was not enough to support operations for two exercises, but in one exercise the ATCOs feedback was that ASR system is adequate for the accomplishment of operations. In this exercise the objective workload measurement (Stroop test) showed a statistically significant decrease (α=0.2%) of workload when ATCOs are supported by ASRU.

ATCOs agreed that the timeliness of information was acceptable, although in some cases they wished it to be higher, and confirmed that the human machine interface, i.e., the ASR output handling was adequate and in one exercise even very satisfactory.

For most exercises the confidence of the ATCOs in the ASR was very high, nevertheless for one was just sufficient.

For some exercises ATCOs considered that the ASR performance was not enough or just sufficient to support them, but for one exercise ATCOs confirmed that level and quality of information provided by the system (in the radar labels) were at an acceptable level.

ATCOs confirmed that ASR did NOT increase the potential for human errors and stated that the situational awareness was increased or unaffected with the introduction of the ASR system.

ASR in the heavy density traffic scenario of the TMA, with high ATCO workload, has been investigated in one exercise of the Solution. The flow could slightly be increase compared to the baseline run (increase of throughput is roughly of one movement per hour in solution runs).

Quality and accuracy of the Solution results are, as it usually happens, a mix of objective measurements and subjective considerations, which are then measured with qualitative assessments using various questionnaire methodologies.

All exercises in the Solution can be considered as having high quality.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

The combination of the joint outcome of the different exercises, provides the results of solution with high significance.

Nevertheless not all the validated concepts have reached TRL6, some are still TRL6 ongoing, so as stated before there was a need of splitting the enabler and create three new ones and have two different definitions on EATMA, one for the part which has reached TRL6 and one for the part which is still TRL6 ongoing. For this purpose a CR has been done.

CR 07562 was created in EATMA to define Solution 96 ASR TRL6 reached as follows:

As most input come from the ATCO-pilot spoken dialog, Automatic Speech Recognition is the appropriate technology to reduce ATCO's workload by directly prefilling command masks from the spoken commands instead of manually inputting them into the system. This requires integration of artificial intelligence (AI) and machine learning algorithms. The solution operates in low to medium TMA and En-Route environment and ATCOs are the end users.

The solution changes the allocation of task between system and human, and the role of the ATCO when entering commands in the system changes from an active role to a monitoring role.

Solution 96 ASR TRL6 ongoing is following defined:

As most input come from the ATCO-pilot spoken dialog, Automatic Speech Recognition is the appropriate technology to reduce ATCO's workload by directly prefilling CPDLC command masks from the spoken commands instead of manually inputting them into the system. This requires integration of artificial intelligence (AI) and machine learning algorithms. Highlighting of targets upon call-sign pronunciation from the ATCO or pilot, user-friendly and intuitive operation will increase controller productivity. The solution will also investigate how ASR may be used to enable faster and more predictable navigation in 3D visualizations of the airspace sectorisation when using dynamic airspace configuration (DAC).

The solution operates in a low to medium En-Route environment and ATCOs are the end users.

6 Recommendations and Additional activities

A set of recommendations have been figured out in order to sharpen ASR operation, supported by AI and Machine Learning, among them:

- Consider a larger amount of representative training data (especially speech data from ATC operations' rooms). Specific training regarding voice recognition and specific ontology and logical rules should be provided to the ASR engine depending on Capability Configuration, Operating Environment and sub-operating environment in which to be deployed, as phraseology varies widely in different Operational environments, such as airport, TMA, continental en-route, oceanic en-route.
- Having separate acoustical models for ATCO and for pilots should be considered as well, as there are differences between their utterances that should be considered for speech recognition, such as differences in pronunciation and structure of utterances and readbacks. Considering the pilot utterances enables reasonable callsign highlighting at ATCo side and readback error detection. There are different opinions on this subject. The HAAWAII project (www.haawaii.de) suggested to use one acoustic model for both. The advance is that this one model can be trained with twice as much of data.
- The command prediction function (ABSR) is not considered necessary for the implementation of Automatic Speech Recognition, but its availability is considered beneficial for event recognition in all situations.
- The basic command prediction function predicting the possible callsigns of the next ATCO or pilot transmissions is, however, a must. At least in one exercise the command recognition rates could be improved by 25% absolute (from 66.9% to 92.1% and also the callsign recognition rate increases also by more than 25% (from71.6%. 97.8%).
- Consider further applications that use the speech recognition and understanding output such as pre-filling of CWP mask and command not sent to the pilot, advanced readback error detection, incident analysis, on-the-job training support.
- Intensify the use and enhance European-wide agreed ontology for annotation of ATC utterances.
- Foster standardization of ASR input and output content as well as format in order to improve system interoperability and comparability.
- Distinguish between automatic speech recognition and automatic speech understanding. Automatic speech recognition just transforms the speech signal into a sequence of words, whereas automatic speech understanding tries to extract the meaning (*descend ten thousand one zero thousand feet* means to descend to 10,000 feet and not to 10,000 plus 10,000 feet).

From the work performed in the Solution a need of standardization has raised. In particular, a standardization of ASR input and output content as well as format would be very useful in order to improve system interoperability and comparability. This concerns speech-to-text with a number of word sequence hypotheses, text-to-concept, i.e. speech understanding, based on the ontology for ATC utterances, and preparations in order to feed succeeding applications such as command error detection, plus formats such as JSON for content transmission, and many aspects more to enable comparability and interoperability. This can be an important topic for future research and collaboration among parties.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

7 Actors impacted by the SESAR Solution

The actors impacted by the Solution are:

- En-Route Controllers, Executive and Planner,
- TMA Executive Controllers,
- Flight Crew.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

8 Impact on Aircraft System

This solution investigates technologies for En-Route and TMA environment and has no impact on Aircraft Systems.

EUROPEAN PARTNERSHIP

Co-funded by the European Union

9 Impact on Ground Systems

This solution has the next impacts on the ground systems:

At a technical system level and to assure good integration, a new resource interaction needs to be deployed between technical systems En-route/Approach ATC and Voice both in ER ACC and APP ACC capability configurations. This is because Voice is in charge of the reception and transmission of audio between the controller and the flight crew. Therefore, a simple voice forwarding resource interaction is deployed between those technical systems: Voice Flow Forwarding.

The following functional blocks and roles have been modified to include new functions:

- Automatic Speech Recognition: encompasses the steps to perform the pure process of speech recognition and the transformation of the sequence of words to instructions by means of the ontology. The speech recognition process may be enhanced based on the input from Flight Planning Lifecycle Management Data Distribution functional block both by command prediction –run on a cyclical basis- and by logical checks performed after the event recognition.
- Controller Human Machine Interaction Management ER/APP: this functional block should include the specific functions related to display and highlight that can be activated by ASR and the prefilling of commands with the information previously recognised, and its presentation for acceptance, rejection or automatic acceptance.
- ATC Executive Controller: this is the role who benefits from the Automatic Speech Recognition. This is the initiator of the process by utterances. In case of display and highlight, the ATC Executive Controller would see the changes in the CHMI and see their situational awareness improved, but no specific functions are required afterwards. In case of the use cases related to prefilling commands, the ATC Executive Controller incorporates functions to check their correctness, and perform acceptance, rejection and correction of the presented commands.

Other relevant functional blocks are:

- Flight Planning Lifecycle Management Data Distribution: this functional block already provides flight plan and sectorisation updates that are now consumed by the ASR so that the latter has access to relevant information regarding the entities able to be affected by speech recognition.
- A/G datalink communications: this functional block is provided for context within the architecture views. It is not supposed to require modifications, as its objective will be composing and sending CPDLC messages as in today's operations at the request of the Flight Planning Lifecycle Management Data Distribution functional block.

10 Regulatory Framework Considerations

The Solution has no impact on the SES and EASA regulatory frameworks.

However, this could change in the future, considering the evolving landscape of the regulatory framework.

11 Standardization Framework Considerations

The Solution has no impact on Standards development, but a need of standardization has raised in the recommendation chapter. In particular a standardization of ASR input and output content as well as format would be very useful in order to improve system interoperability and comparability. This concerns 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 command error detection, plus formats such as JSON for content transmission, and many aspects more to enable comparability and interoperability. This can be an important topic for future research and collaboration among parties.

ASR should be able to interact with VCS under standard ED-137 from EUROCAE.

Input from ground system to ASR is covered by ED-137 and ASTERIX-62, whereas input from ASR into ground system would benefit from a dedicated new standard in order to be easily integrated in any ATM system. This topic deserves further investigation with the involvement of EUROCAE in order to state the actual need of an output standard and following defining the appropriate one.

12 Solution Data pack

The D4.1-PJ.10-W2-96 ASR solution pack TLR6 includes the following documents:

<table>
<tr><th colspan="2">Solution Data Pack</th></tr>
<tr><td rowspan="4">Systems Requirements consolidation</td><td>Technical Specification/Interface Requirements Specification (TS/IRS) TRL6 - Final version
D.4.1.020
Edition date: 21/03/2023
Edition: 01.00.00</td></tr>
<tr><td>TS/IRS Part II (SAR)
D.4.1.020
Edition date: 22/03/2023
Edition: 01.00.00</td></tr>
<tr><td>TS/IRS Part IV (HPAR)
D.4.1.020
Edition date: 21/03/2023
Edition: 00.02.00</td></tr>
<tr><td>PAR
(Additional document for technological solution, non PMP)
Edition date: 06/03/2023
Edition: 00.01.00</td></tr>
<tr><td>Cost Benefit Analysis tailored for the specific Technological Solution (CBAT) TRL6</td><td>Cost Benefit Analysis (CBAT) – TRL6
D4.1.110
Edition date: 23/03/2023
Edition: 00.01.01</td></tr>
<tr><td>Technical Validation Report (TVALR) TRL6</td><td>Technical Validation Report (TVALR) – TRL6
D.4.1.100
Edition date: 30/03/2023
Edition: 00.02.00</td></tr>
</table>