SafeTeam

D3.2 Human Factors Design Principles for a Stabilised Approach Digital Assistant

Document number	D3.2
Document title	D3.2 Human factors design principles for a stabilised approach digital assistant
Version	1.0
Work package	WP 3
Edition date	29.06.2023
Responsible unit	Technical University of Munich
Dissemination level	PU
Project acronym	SafeTeam
Grant	101069877
Call	Safe, Resilient Transport and Smart Mobility services for passengers and goods (HORIZON-CL5-2021-D6-01)
Topic	HORIZON-CL5-2021-D6-01-13: Safe automation and human factors in aviation – intelligent integration and assistance

This project has been funded by the European Union under Grant Agreement 101069877

Funded by the European Union

SafeTeam Consortium

innaxis	Innaxis (INX)
AESA	Agencia Estatal de Seguridad Aérea (AESA)
TUM	Technische Universität München (TUM)
DataBeacon	DataBeacon
ONERA THE FRENCH AEROSPACE LAB	ONERA
RI. SE	Rise Research Institutes of Sweden AB (RISE)
PEGASUS AIRLINES	PEGASUS HAVA TASIMACILIGI ANONIM SIRKETI (PEGASUS)
UK Civil Aviation Authority International	CAA INTERNATIONAL LIMITED (CAAi)
L3HARRIS™	L3HARRIS FLIGHT DATA SERVICESLIMITED

Document change record

Version	Date	Status
1.0	2	Initial Submission

Abstract

In work package 3.2 of the SafeTeam project, human factors are investigated to safely integrate AI into the workflow of modern-day airliners. More specifically, we aim to develop an unstable approach prediction tool to increase the pilots' situational awareness during the critical phases of an approach. This document represents the first of two reports that document the development process of such a tool. This first report conducts a thorough literature review and describes the methods we intend to use. Further, the use case is defined, and several models are developed to gain an in-depth understanding of the current system and how changes to this system may impact positively and negatively the work in a cockpit during the approach. The work presented here is the foundation of developing an AI decision support tool that will be implemented and tested in a simulator study later in the project.

Funded by the European Union

Table of Contents

Funded by
the European Union

List of Tables

List of Figures

Funded by
the European Union

Introduction

In the SafeTeam project, human factors related to AI integration are elaborated through several practical applications, in which we aim to progress the safe introduction of automation in the form of intelligent assistance to humans. In the present deliverable, we lay the foundation for developing a stabilized approach assistant, based on a machine learning unstable approach prediction tool, that is intended to increase the pilot's situational awareness during the approach phase of a flight.

A share of 43% of fatal accidents occur during approach and landing, while these flight phases account for only 16% of total flight time [1]. In the years 2015 to 2019, IATA lists 292 flight accidents for turbine-powered aircraft with a certificated Maximum Takeoff Weight (MTOW) of at least 5700 kg. In 17 % of these accidents, the IATA ACTG identified a UA as a contributing factor. In the major accident category, namely runway and taxiway excursions, UAs were identified as contributing to 22 % of absolute occurrences [2]

To identify risks and avoid accidents before they occur, operators' flight safety teams routinely analyze flight data recordings stored in the onboard Quick Access Recorder (QAR). Recorded flight parameters include accelerations, speeds, altitudes, coordinates, attitudes, configurations, control surface deflections, system statuses, and many more. Flight data analysis, also called Flight Data Monitoring (FDM) or Flight Operations Quality Assurance (FOQA), aims to detect exceedance events in examined recordings. Exceedance events occur when flight parameters exceed their corresponding predefined thresholds such that the safe operation of a flight is not maintained [3].

Based on QAR data, several machine-learning applications have been developed to predict unstable approaches [4] [5] [6]. Their actual implementation into the operational environment ,however is not covered by literature to this extent.

This deliverable provides a case study definition that will enable a practical case study for the stabilized approach digital assistant, one of the three investigated case studies for the safe teaming of humans and Ai in aviation. The focus of this document is the documentation of the following aspects:

- Operational: Expected use (by who, when, how), user challenges to overcome (environmental, performance, under-capacity, safety)
- Technological: Current maturity level, data availability, targeted level of autonomy, AI performance concepts (e.g., explainability, accuracy, recall, and bias-variance trade-off)
- Human factors: distribution of tasks and responsibilities among operators and technology, human-machine interface upgrades for safe implementation, human technology safe teaming to reduce workload.

Based on the results of SafeClouds.eu, which demonstrated the technical feasibility of machine-learning-based unstable approach predictions, this document will guide the necessary developments that need to be implemented in work package 4.2, the Human-machine Collaboration in Destabilized Approaches. Therefore, this document develops:

- a list of required functionalities for this digital assistant by incorporating the requirements of users and relevant stakeholders, with the goal of evolving this assistant as an automatized detector/predictor of unstable approaches for the cockpit
- a use case definition by identifying novel data sources that could improve its precision, along with a data collection plan that studies the availability of those data sets.
- an evaluation of how well the available AI models for go-around predictions fulfill user requirements in operational use.

Funded by the European Union

- human factors-based design of the assistant interface to present the information to the crew in a safe manner, avoiding information overload and defining the appropriate task distribution between humans and machines according to the envisioned autonomy level.

The structure of the deliverable is as follows. First, we provide a literature overview of existing machine-learning-based unstable approach predictions, which is followed up by a description of the unstable-approach prediction developed in the H2020 project SafeClouds.eu. This model serves as the basis for developing an initial Concept of Operations for the stabilized approach digital assistant case study. In section 2, we summarize the methodology applied for developing the initial Concept of Operation, a combination of parts of the SafeTEAM Framework, which will be presented in detail in Deliverable 2.1, and parts of EASA's guidance on machine learning applications in aviation. Finally, the Use Case definition is presented in section 3, which includes the following:

- a description of the approach phase relevant to the case study
- a description of how to integrate the machine learning-based component in the approach phase
- an initial set of requirements that will guide the implementation phase of task 4.2.

1 Literature Review

Before defining the use case and case study for task 3.2, we provide an overview of the work that has been done in machine learning regarding the prediction of unstable approaches. Therefore, we first present results from a literature review and follow up, with a summary of the relevant results from the SafeClouds.eu project, which serves as an asset for the SafeTEAM project.

1.1 Review of Machine Learning-Based Unstable Approach Predictions

Over the past decade, some research has been performed on predicting unstable approaches or go-arounds. Machine-learning-based unstable approach prediction is at the heart of this case study's concept. A brief overview of the existing research is provided in this section. The integration of the unstable approach predictors into the operations needs to be covered more in the literature.

Early work on machine-learning-based unstable approach predictors was performed in the work of Wang [4]. The motivation to study Unstable Approach Predictors is to provide lead time to the flight crew to adjust an approach and avoid potentially unstable approaches and consequently go-arounds as one significant knock-on effect. Based on 28 days of surveillance flight track data for approaches on Newark International Airport Runway 22L, a logistic regression model is trained to nowcast the likelihood of the approach stability, given the flight's performance situation at 10NM, 6NM, and 3NM from the runway threshold. The approach stability is evaluated for two intervals, first between 1000ft AGL and runway threshold and between 500ft AGL and runway threshold.

The training and test data set was generated by splitting the overall 8158 approaches randomly into a training set containing 5000 approaches and a test set having 3185 approaches.

According to Wang, 52.4% of the approaches satisfy the stabilization criteria from the 1000ft AGL to the runway threshold interval, and 82.3% meet the stabilization criteria from the 500ft AGL to the runway threshold. Compared with studies from the Flight Safety Foundation [7], which state an Unstable Approach rate of 3-5%, these numbers appear too high and raise questions about the applied stable approach criteria used in the labeling process of the data set or the missing imbalance of (un-) stable approaches in the data.

In the initial publication [4], ten features were derived from the data set, including, e.g., Maximum Take-Off Weight (MTOW), groundspeed, descent rate, or deviation from the runway centerline. In a subsequent publication [8], the feature set was enlarged to 22 features, including weather information. While claims to improve the performance measures which is true, however, in only small gains, the reported performances and the ones used for comparisons could be more consistent, making a final judgment on the reported results difficult. The achieved precisions are summarized in

Table 1 as follows:

Table 1: ML Precision Results from [4]

Prediction Location from Runway Threshold	1000ft AGL Interval	500ft AGL Interval
10NM	62,7%	40,0%
6NM	76,0%	52,6%
3NM	86,7%	53,0%

Funded by the European Union

The achieved recall rates from [4] are summarized in Table 2 as follows:

Table 2: ML Recall Results from [4]

Prediction Location from Runway Threshold	1000ft AGL Interval	500ft AGL Interval
10NM	45,8%	3,6%
6NM	63,9%	16,5%
3NM	76,2%	31,4%

The conclusion stated, "This analysis shows the degree of accuracy that can be achieved using historical flight track data to nowcast stabilized approaches before reaching 1000' AGL and 500' AGL. Whether this is useful on the flight deck and how this "probabilistic" information could be integrated into "deterministic" flight deck procedures are open research questions." The presented results are an initial step in predicting unstable approaches. Due to the inconsistency between the two mentioned publications and the not imbalanced data set, the presented results are considered not sufficiently conclusive for this project.

In [5], unstable approach predictions were tested based on a set of 79766 approaches from 79667 recorded flights from Quick Acess Recorder data. The work is based on a detailed unstable approach labeling algorithm, an evolution of the algorithm developed in SafeClouds.eu. It takes into account the following:

- speed difference between calibrated airspeed and target speed
- altitude rate
- ILS deviations
- Flap Position
- Gear Position
- Power Setting

with a decision height fixed at 1000 ft above runway threshold elevation, coming with the limiting assumption of only considering ILS approaches. The results of the labeling can be summarized as follows:

- 3115 unstable approaches were detected in the data (3.91%)
- 215 approaches performed were go-around (0.27%)
- 97.43% of unstable approaches did not perform a go-around
- 36.04% of go-arounds were detected to be unstable approaches

which are well aligned with the numbers published by the flight safety foundation [7].

A set of features was extracted from the described data set for each approach. In a time interval of 45 seconds before reaching the stabilization gate, the features were extracted every 5 seconds. For a time interval 80s before the time reaching the stabilization gate, standard deviations of the features indicated with a * were additionally extracted.

Featire Categpry	Features
Aircraft Handling	Standardized Speed Difference, Barometric Altitude (ARTE), Glideslope Deviation*, Localizer Deviation*, N1 Rotational Speed*, Ground Speed, Calibrated Airspeed, Altitude Rate*, Temporal Margin Speed, Difference Aircraft Track and Runway Bearing*

Aircraft Configuration	Flap Handle to Landing, Gear Position, Temporal Margin Flap Extension, Temporal Margin Flaps in Landing Configuration, Temporal Margin Gear extended
Flight Static Information	Time Flown, Captain Flying, Aircraft Mass, Aircraft Type, Number of previous GAs
Weather On-Board	ADIRS HWC, ADIRS CWC
Weather METAR	Wind Speed, Gusts, Visibility, Temperature, Dew Point, Local Pressure, Weather Phenomena

A benchmark study for twelve predictors was performed based on the labeled feature data set. All predictors were trained to predict the stability of an approach 20 seconds before the stabilization gate. The best-performing model's result is illustrated in Figure 1. The precision-recall area under the curve (AUC) metric achieved is 0.608.

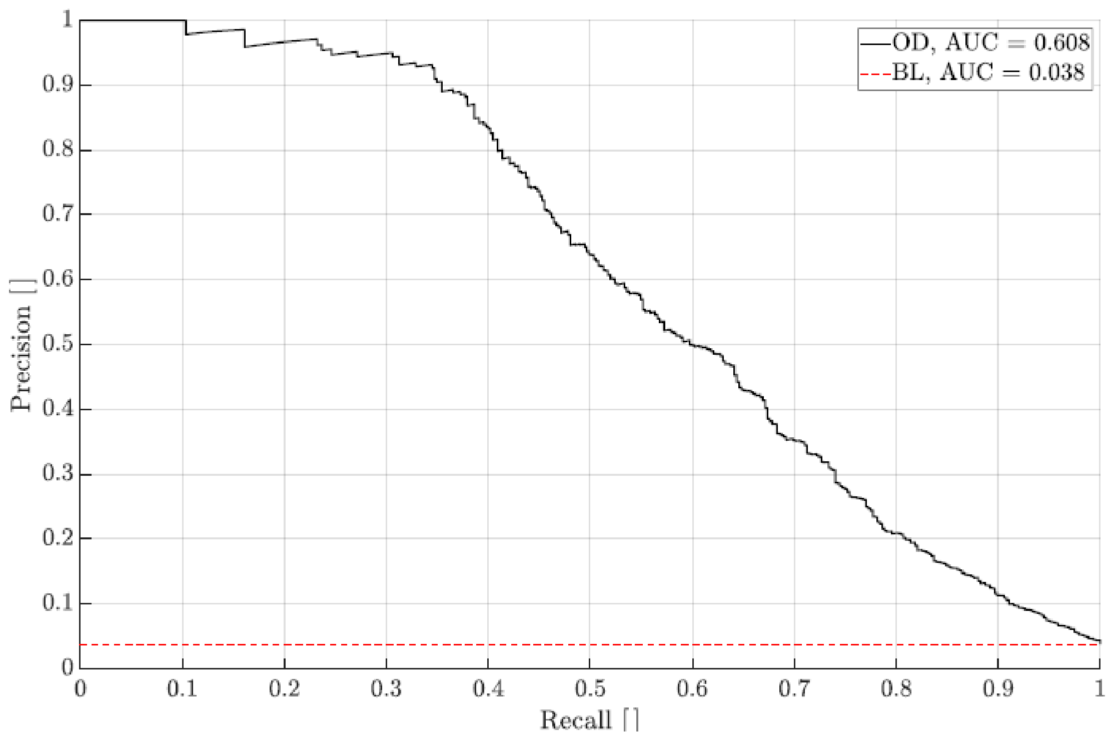


Figure 1: ML Results from [5], summarized in a Precision Recall Curve

Figure 2 illustrates the precision recall evaluation for different prediction horizons, from 5s to 40s. One can see, that with larger prediction horizons, the AUC value decreases, which meets the intuitive expectations.

Funded by the European Union

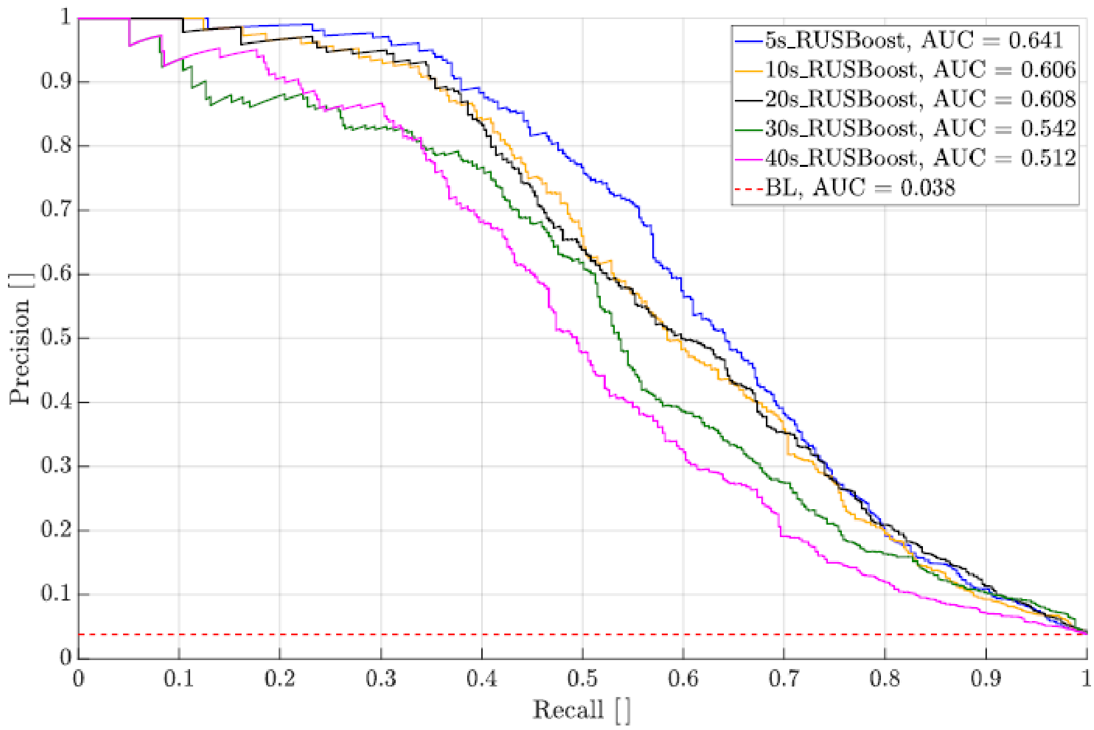


Figure 2: Precision Recall Results from [5], for different prediction horizons

Another interesting work has been done [9]. As this work covers Unstable Approaches in general aviation (GA), the labeling process differs from the works covering commercial aviation. The paper is motivated by the observation that unstable approaches pose a significant risk to general aviation. Therefore, the paper investigates a deep learning architecture called recurrent neural networks (RNNs) to develop a predictive warning system to reduce risks associated with UA in GA.

The paper presents a data pipeline for labeling and feature engineering. For altitudes between 450 ft mean sea level (MSL) and 300 ft MSL, the following criteria were used:

- an indicated airspeed of 79–94 kt,
- a vertical speed of less than 1000 ft/min,
- pitch of −10 to 10 deg,
- bank of −10 to 10 deg
- engine 1 and engine 2 have a percentage power of less than 40%.

For altitudes between 300 ft MSL and 150 ft MSL, the following criteria were used:

- an indicated airspeed of 79–89 kt,
- a vertical speed of less than 750 ft/min,
- pitch of −6 to 6 deg,
- bank of−7 to 7 deg,
- engine 1 and engine 2 have a percentage power of less than 35%

Based on these criteria, a data set containing 22,512 unstable labeled approaches and 19,502 labeled as stable. As features used to train the ML solution, the following information is extracted from the raw data:

- altitude in ft above MSL,
- calibrated airspeed in kt,

- vertical speed in ft/min,
- pitch angle in degrees,
- bank angle in degrees,
- engine 1/2 power in %.

The achieved model performance is provided in Table 3. In the paper, only the precision and recall regarding the Unstable Approach Prediction are presented; the values for the SA class are computed from the numbers given.

Table 3: Summary of ML Results from [9]

Class	Number of Approaches	Precision	Recall
SA	19502	0.9007	0.7647
UA	22512	0.7895	0.9128

The share of unstable approaches is higher than 50%, which looks high, even for general aviation. Therefore, the problem of imbalanced data, especially encountered and addressed [5], must be more relevant for the predictor discussed in this paper. Also, the set of features is limited and only considers aircraft performance-related information but does not contain weather or airport-related information. The altitudes for the labeling are defined against MSL, which needs to be clarified for us since we would expect a definition against the runway threshold elevation (ARTE). Since Embry-Riddle Aeronautical University is involved, all approaches may have been performed at Dayton Beach Airport, which is close to the sea. Thus the MSL and ARTE can be considered identical. Overall the data set is more comparable to an ADS-B-generated data set than a data set generated from a commercial aviation FDM program, even though the temporal resolution is stated to be one second, which is considerably better than the average ADS-B temporal resolution.

1.2 SafeClouds.eu Unstable Approach Prediction Summary

In SafeClouds.eu, different case studies investigated data mining techniques to obtain predictive information on safety events [10]. The case study "Unstable approaches" (UA) investigated two machine learning solutions [6]. The study in SafeCloud.eu focused primarily on the technical feasibility and the underlying IT infrastructure necessary to develop the machine learning models. As these models serve as a starting point for further developing the stabilised approach case study, we present a detailed summary of the SafeClouds.eu results.

For this, a series of data, mainly FDM (Flight Data Monitoring), were used, which can be seen in Table 4

Table 4: Data Sources from SafeClouds.eu

Data category	Source	Data type	Description
FDM	Airlines consortium member	Binary file decoded into engineering values and stored in parquet format	Contains static and dynamic information as recorded in the QAR device, such as barometric altitude, ground speed, radio altitude, etc. with meta information (e.g. airport ICAO codes).
METAR	Iowa State University	Text files converted into parquet file	Contains information about the weather situation at an airport. It includes wind, temperature, visibility, cloud ceilings, precipitation, etc.

Funded by the European Union

Airport related information	OurAirports website	CSV files converted into parquet files for performance	Contains airport information such as the airport ICAO code, its region, country, position, runways, navigation aids, etc.

For the identification and labelling of these UA, a standard definition was developed by all the consortium stakeholders, establishing the limits and requirements of different flight variables to flag an approach as unstable or not. Table 5 explains the variables/requirements defined to be used in the labelling of UA. The implementation considered varying limits for approach speeds and ILS deviations (localizer and glide slope) based on different height intervals. The final criteria used for the labelling of UA can be found in Table 6. These criteria established three limits by severity (limHI, limMD, limLO) within a certain altitude range (altBeg_ft, altEnd_ft) for a specific approach type (IMC, VMC, CIR) that should be achieved for a minimum duration (tHI, tMD, tLO). Ultimately, UA were classified as high, medium, or low severity. The highest severity would be assigned if an UA obtains several severity levels.

Table 5: Naming convention of UA Criteria

ID	Description	ID	Description
vAppU1 [kts]	Upper limit on approach speed deviation for first height interval	spBr [ft]	height limits for the speed brake usage
vAppU2 [kts]	Upper limit on approach speed deviation for second height interval	gear [ft]	height limits for gear extension
vAppL1 [kts]	Lower limit on approach speed deviation for first height interval	flap [ft]	height limit for final flap setting
vAppL2 [kts]	Lower limit on approach speed deviation for second height interval	bAng1 [deg]	limit for bank angle for first height interval
vertSp [ft/min]	Limit on the vertical speed	bAng2 [deg]	limit for bank angle for second height interval
loc [dot]	limit on localizer deviation	fanSp [%]	limit on minimum fan speed of the engine
gs [dot]	limit on glide slope deviation		

Table 6: Limits for unstable approach labeling

	ID	IMC	VMC	CIR	altBeg_ft	altEnd_ft	limHI	limMD	limLO	tHI	tMD	tLO	deltaLim
0	vAppU1	True	False	False	1000	500	30.0	20.0	15.0	3	3	3	NaN
1	vAppU2	True	True	True	500	50	20.0	15.0	10.0	3	3	3	NaN
2	vAppL1	True	False	False	1000	500	-10.0	-5.0	-2.0	3	3	3	NaN
3	vAppL2	True	True	True	500	50	-5.0	-3.0	-1.0	3	3	3	NaN
4	vertSp	True	True	True	1000	50	-1200.0	-1100.0	-1000.0	3	3	3	NaN
5	spBr	True	True	True	2000	50	1000.0	1250.0	1500.0	0	0	0	500.0
6	gear	True	True	True	2000	50	1000.0	1250.0	1500.0	0	0	0	500.0
7	flap	True	True	False	2000	50	700.0	1000.0	1200.0	0	0	0	500.0
8	flap	False	False	True	1000	50	300.0	500.0	700.0	0	0	0	300.0
9	bAng1	True	False	False	1000	500	20.0	17.0	15.0	0	0	0	NaN
10	bAng2	True	True	False	500	50	15.0	12.0	10.0	0	0	0	NaN
11	bAng2	False	False	True	300	50	15.0	12.0	10.0	0	0	0	NaN

12	loc	True	True	False	1000	50	1.5	1.2	1.0	3	3	3	NaN
13	loc	False	False	True	300	50	1.5	1.2	1.0	3	3	3	NaN
14	gs	True	True	False	1000	50	1.5	1.2	1.0	3	3	3	NaN
15	gs	False	False	True	300	50	1.5	1.2	1.0	3	3	3	NaN
16	fanSp	True	True	True	1000	50	NaN	NaN	NaN	0	0	0	10.0

Figure 3 illustrates the final workflow in the UA labeling. Before this, preliminary data processing was carried out, including actions barometric altitude correction, time point determination of the selected altitudes (e.g., 500ft and 1000ft), and approach information such as approach type, METAR information, approach speed, and the runway identifier. The vertical speed is first calculated using precomputed approach speed and the targeted runway's glide slope during the labeling process. Next, relevant indicators and limits are selected based on the approach type. ILS data (LOC and GS) validity is verified before entering the main loop for all indicators. The label functions for each indicator share the same interface.

Funded by
the European Union

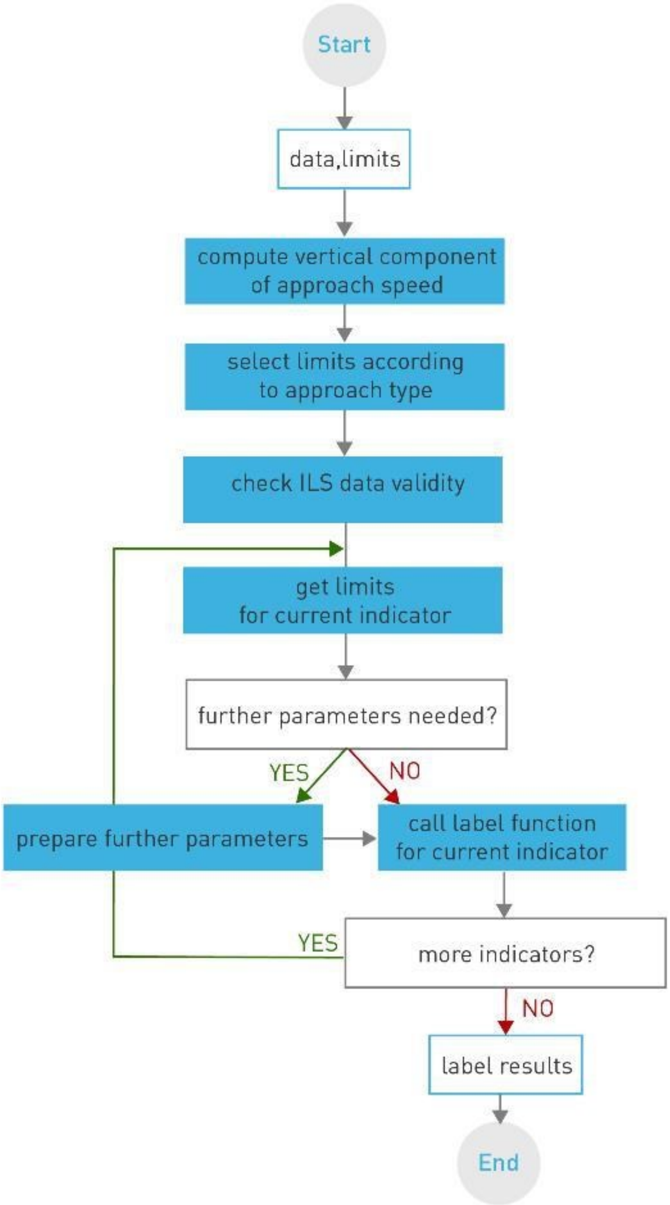


Figure 3: Top-level control flow of the labeling module

Figure 4 shows an overview of the feature engineering process. In this case the features for predictive learning were computed during the same stage as the labeling of UA events mainly due to the reduction of computational cost. The feature extraction works by calculating a list of selected features at each time point of interest. This includes current measurements and those taken over short intervals of 60 seconds. Static features are also calculated and merged, resulting in approximately 600 numerical values. It is essential to highlight that each landing attempt is handled separately.

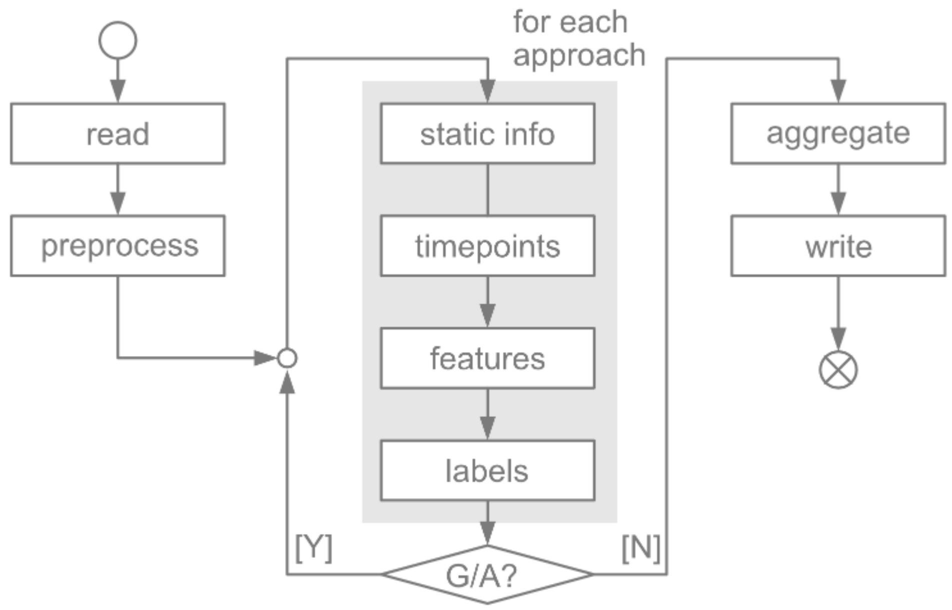


Figure 4: Data processing and feature extraction pipeline

Table 7 illustrates the primary defined features of the UA case study. They have been grouped according to their nature. All features are computed for each of the specified time points of interest. Features whose name ends in "_var" refer to calculating the variance of the relevant samples in the last 60 seconds. Other features such as "airspeed_mDs" (aircraft airspeed in meters per second) represent a static measurement at that very moment. METAR information and the "number_of_holdings" are supplied once for the whole approach since they will not change as frequently. For the "handling qualities" group of features, the aim was to establish how calm/turbulent the aircraft behaves in the approach phase based on the variation of altitude, heading, pitch, or roll values, among others. The "energy_level" feature is derived using the barometric altitude of the aircraft. For "adverse weather" features, wind direction, speed, and variation were also computed using the METAR weather reports. The aircraft control indicators in FDM were analyzed to determine the pilot who was flying. For example, if the left autopilot is on, and the left sidestick (in Airbus aircraft) or right transmitter is active, it's most likely the captain. If it's the opposite, it's likely the first officer. The pilot awareness features were defined through the duration and distance of the flight and the number of holding patterns performed. To calculate the number of holdings, we add up all turns in either direction and determine if the sum is divisible by 360° within a 5-minute tolerance (the standard holding pattern is 4 minutes, but some margin of error was added). Finally, "surrounding traffic" features at the target airport were defined due to the possible effects of the pressure/workload on the pilots and their approach quality. The density of air traffic control (ATC) communications via VHF activity was recorded to define this. For a more detailed examination of the approach situation, the radar track data from ALLFT+ shortly before landing was analyzed, focusing on the position and speed of the leading and trailing aircraft and measuring the distance and speed differences. For a more detailed definition of the features and the feature extraction process, see SafeClouds deliverables D4.2 [10] and D4.3 [6].

Table 7: Primary Features, computed from the SafeClouds.eu Data Pipeline

Feature Group	Features
handling quality	pitch_rad_var, roll_rad_var, heading_rad_var, aoa_rad_var, p_radDs_var, q_radDs_var, r_radDs_var
aircraft energy	airspeed_mDs, energy_level, gndspeed_mDs, hbaro_m, hdot_mDs, mass_kg, rheight_m

Funded by
the European Union

adverse weather	pstatic_NDm2, wind_dir_rad, wind_spd_mDs, wind_dir_rad_var, wind_spd_mDs_var, METAR (static)
configuration	flaps_rad
crew coordination	pilot_flying (includes autopilot status)
pilot awareness	distance_m, flight_time_s, utc_time_s, number_of_holdings
surrounding traffic	vhf_keying_var, airport throughput (ALLFT+)

Looking into the data, it was observed that instability during the approach typically occurred around three nautical miles (NM) before the runway threshold. It was decided that for the prediction to be useful to the pilots, it should give them a range of reaction time between 30-90 seconds. To account for this, the prediction point needed to be established between 4 and 9 NM from the threshold, giving the pilot enough time to react. At a minimum, the pilot has 1 NM, or approximately 30 seconds of travel time, to prevent a UA from occurring.

One of the main challenges encountered, and one to highlight, during the development of data-driven prediction tools is the imbalance in the data. UA incidents are rare events, occurring in only about 5% of all approaches. This leads to some issues in correctly predicting actual UA incidents, resulting in many cases of a high false positives rate. To address this issue, synthetic samples were generated using a technique known as SMOTE (Synthetic Minority Oversampling Technique) to support the model in learning features and address the lack of positively labeled flights (UA events).

In the end, two different approaches were taken to predict UA events. On the one hand, **Prediction case A** used machine learning to perform precursor analysis (feature selection and combinations) and provide a binary classification model for predicting a flight's stability with enough time for the pilot to stabilize. On the other hand, **Prediction Case B** used deep learning techniques applied to the flight data management (FDM) time series using neural networks for binary classification of an impending unstable approach and to detect unseen hazards or anomalies in approach procedures.

For prediction case A, one of the main objectives was to achieve maximum accuracy in predicting UA flights while minimizing false negatives, ensuring that the model's predictions do not result in false alarms and positively impact air traffic operations and increase throughput. The prediction point was established at 4 NM from the runway threshold, and features were sampled every 0.5 Nm from 4 to 9 NM. An initial comparative study was conducted to test the initial performance of 7 different ML models. The ensemble-type model called LightGBM was obtained from these results as the most suitable model. This model was then re-trained and fine-tuned. The final performance metrics can be seen in

Table 8.

Table 8: Prediction case A: Model evaluation metrics

Class	Precision	Recall	Specificity	F1-score
Not UA	0.97	1.00	0.53	0.98
UA	0.85	0.53	0.99	0.65
Avg / Total	0.97	0.97	0.56	0.9
AUC (ROC)	0.96			
AUC (PR)	0.77			

The model's results showed that it could accurately predict regular non-UA flights. However, the limited number of samples for UA flights led to it having some issues detecting certain types of UA. The model had a high precision score for the negative class (UA), providing a suitable prediction with high confidence levels of 1NM before the event. Despite this, the model was selective, having low recall due to the limited number of negative samples and, thus, difficulty recognizing specific unstable approaches. In addition to the predictions, this type of ML model also allowed for an analysis of the feature importance determined by the model. This allows us to look and try to understand why the model makes the predictions it does. Figure 5 presents feature importance, sorted by their impact on model prediction. The top feature identified by the model was "weather_altitude_hpa," representing the destination airport's QNH. This is followed by airspeed at 4 NM, barometric altitude and airspeed with fully deployed flaps, and aircraft descent height over time at 4NM. Mainly features around 4NM to 5NM presented a higher significance level for UA prediction.

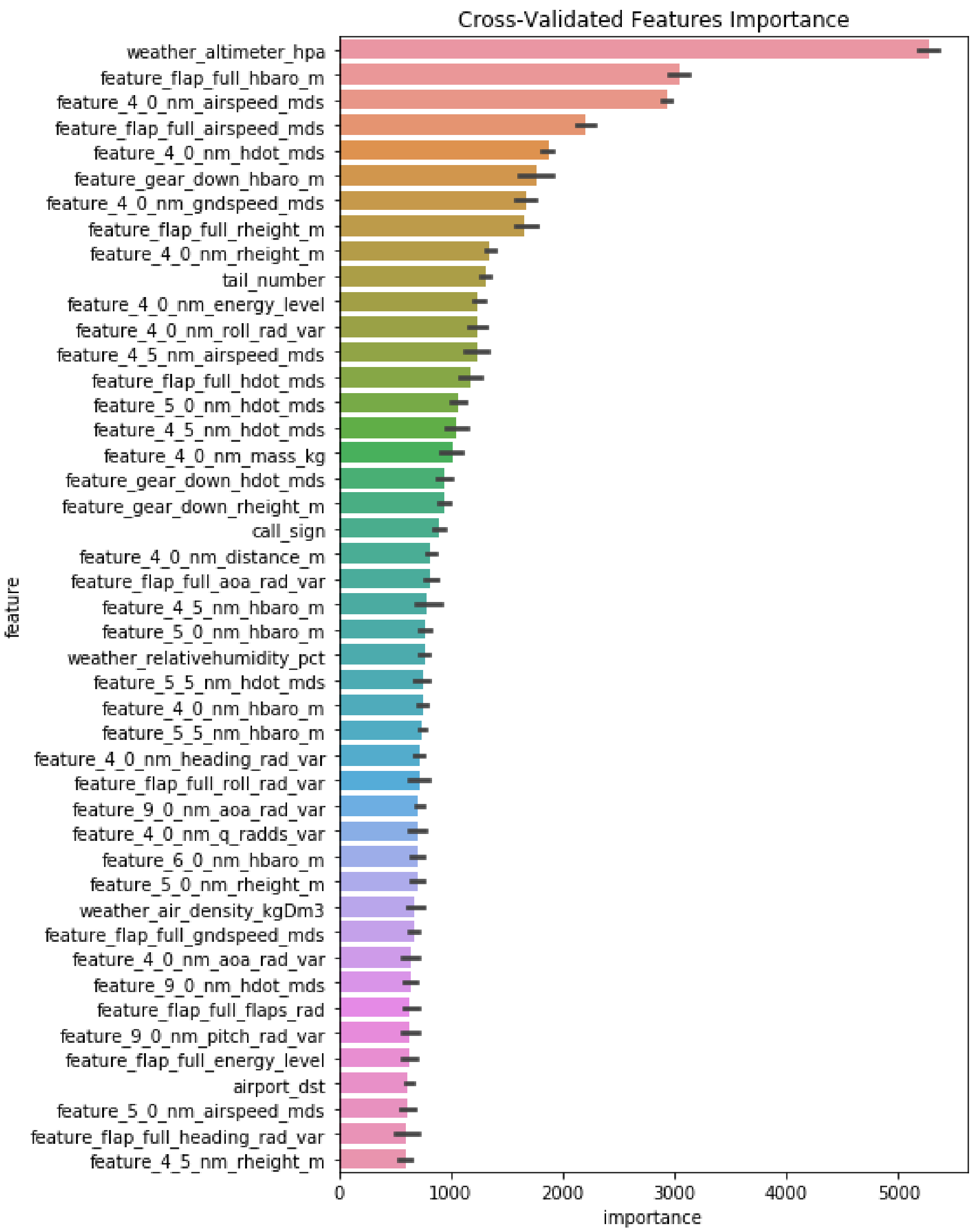


Figure 5: Prediction Case A Feature Importance

Funded by the European Union

For prediction case B the main objectives was the same: to achieve maximum accuracy in predicting UA flights while minimizing false negatives, ensuring that the model's predictions do not result in false alarms and positively impacting air traffic operations and increasing throughput. However, in this case the prediction would be given to a time series observation. This means that instead of having a static prediction point as in case A, we now have a dynamic prediction point. The condition set for this case study was that the final ML model developed should be able to provide a correct UA prediction at least 30 to 60 seconds prior to the event. A LSTM (Long short-term memory) neural network was used which is a type of recurrent neural network (RNN) mainly used on sequential or time series data. These kinds of models are capable of automatically extracting features from past events and LSTMs are specifically known for their ability to extract both long and short term features. The final performance metrics for the trained LSTM can be seen in Table 9.

Table 9: Prediction Case B: Model Evaluation Metrics

Class	Precision	Recall	Specificity	F1-score
SA	1.00	0.94	0.97	0
UA	0.77	0.98	0.87	1
Avg / Total	0.97	0.97	0.56	0.9
micro AVG	0.95	0.95	0.95	micro AVG
macro AVG	0.88	0.97	0.92	macro AVG
weighted AVG	0.96	0.95	0.95	weighted AVG

The model performance showed a low precision level implying that some non-UA observations in the approach phase are being identified as potentially Unstable Approach (UA). One reason could be that these are situations in which these approaches present very similar conditions to UA ones but may not result in a UA as the pilot may have taken corrective action within the seconds between the prediction point and the actual UA (reaction offset). On the other hand, the recall metric shows high reliability in predicting a UA, correctly identifying most of them, indicating that the neural network effectively learned the conditions preceding a UA. For this case study, the feature importance was also analyzed, highlighting as main features the calibrated airspeed, the ground speed, the relative humidity from METAR, the flaps position, and the barometric altitude. Figure 6 presents features important for case B.

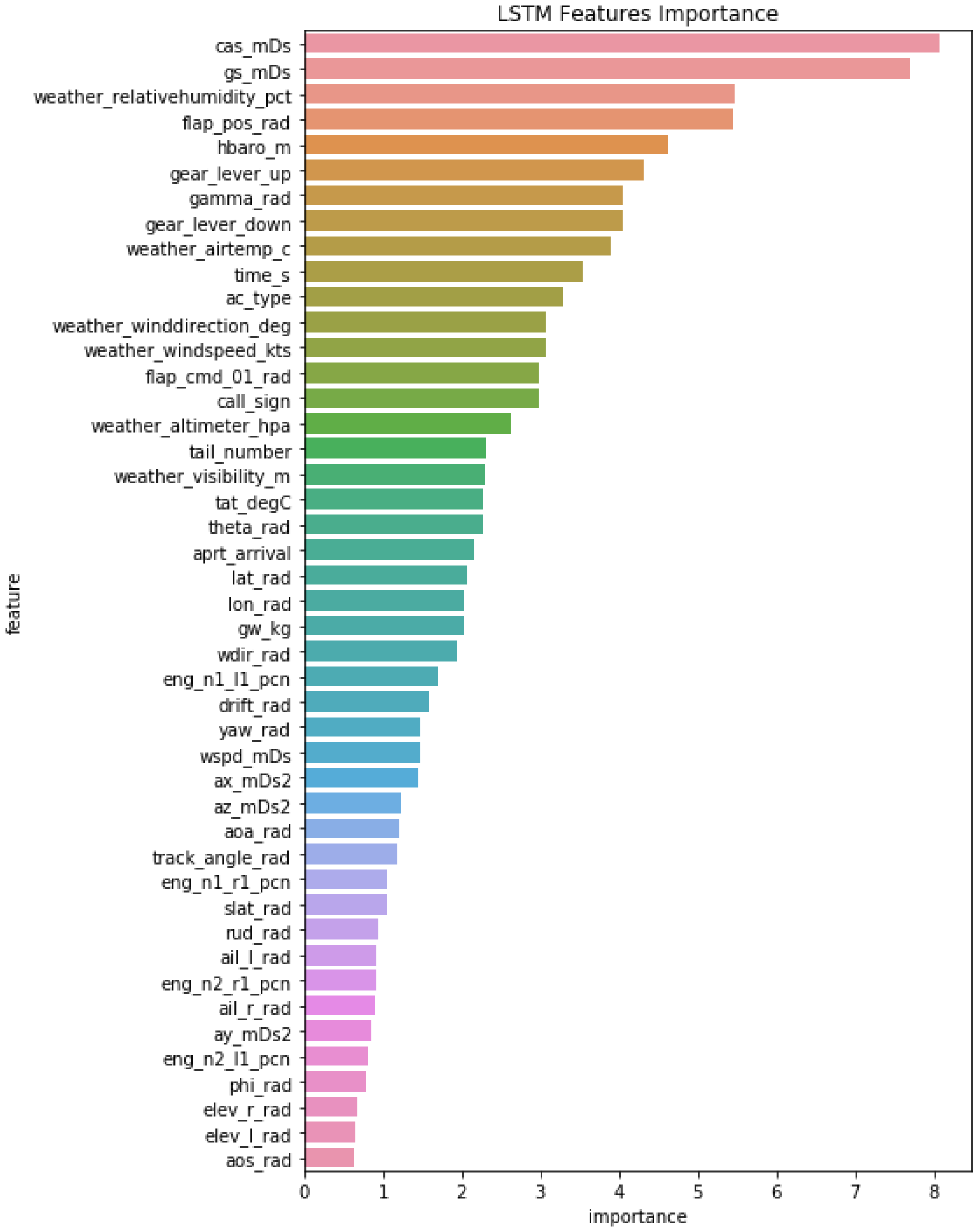


Figure 6: Feature Importance Prediction Case B

One key lesson learned from the Safeclouds project is the importance of explainability in the context of safety applications. The significance of comprehending machine learning models' inner workings to cultivate trust and acceptance among end-users, particularly within safety-critical domains, was emphasized. A preliminary assessment of explainability was carried out using techniques such as feature importance analysis and partial dependence plots. While this initial assessment provided valuable information, it is clear that current practices could further improve the explainability of the models. The field of explainable artificial intelligence (XAI) has evolved rapidly, offering innovative methods such as rule extraction, counterfactual explanations, and model diagnostic techniques such as LIME and SHAP. Some of these techniques have already been successfully applied in other projects such as SafeOPS [11], [12]. Incorporating these advanced techniques in future iterations of the ML models would deepen our understanding of the predictive factors behind unstable approaches, provide us with the opportunity to improve the pilot's situational awareness by providing more decision-useful information, and at the same time, reinforce the pilots' trust in the system's decision-making process.

Funded by the European Union

Furthermore, another key lesson learned was the need from the beginning of the project to establish a clear definition of the problem and to establish through the collaboration of users and developers a set of valuable and realistic requirements for the ML model performance. Safety-critical applications like predicting unstable approaches may require overall accurate predictions and a thorough understanding of the precision-recall trade-off. The precision-recall tradeoff is crucial in AI/ML applications. Precision measures accurately identified positive instances, while recall gauges the ability to detect all positive instances. Achieving an optimal balance between precision and recall is crucial to uphold the reliability and safety of the system. This delicate equilibrium necessitates careful attention during the training process. Depending on specific requirements, the training can be optimized to prioritize either precision or recall. For example, if the objective is to minimize false alarms, placing a higher emphasis on precision is favorable. Conversely, if the aim is to prevent the omission of critical instances, a greater focus on recall is prioritized. Accomplishing the desired balance entails meticulous parameter tuning, metric selection, and adjustment of decision thresholds.

Finally, it is worth mentioning that the SafeClouds project achieved a Technology Readiness Level (TRL) 4: technology validated in the lab. TRL4 is described as following the successful "proof-of-concept" work: basic technological elements must be integrated to establish that the different parts will work together to achieve concept-enabling performance levels for a component and/or breadboard. This validation must be developed to support the previously formulated concept and should also be consistent with the requirements of potential system applications. The validation could be composed of ad hoc discrete components in a laboratory. On the one hand, laboratory validation was carried out to validate the FDM decoding and assess whether the errors found in the case studies were due to the recording systems or our own decoding configuration. For this purpose, samples decoded by the airlines were used. On the other hand, validation was carried out to assess if labeling flights as (un)stable followed the designed rules and that the designed rules corresponded to the airlines' best practices. To this end, we verified that the criteria were correctly applied and identified possible risks in the process and their mitigations.

1.3 Cockpit Design Principles

Cockpit design philosophy and HMI best practice principles were gathered and customized for use in the stabilized approach digital assistant design. Sources for this information include literature as [13], [14], and [15], discussions with leading experts in human factors and cockpit design, information presented in robotics system design classes, and discussions with professional airline pilots.

Cockpit design philosophy and HMI best practice principles were gathered and customized for use in the stabilized approach digital assistant design. Sources for this information include discussions with leading experts in human factors and cockpit design, information presented in relevant lecture classes, various research publications, and discussions with professional airline pilots.

These statements establish a basis for system expectations, which will be developed later into system requirements.

Classic HMI Design Principles

1. Clear and Intuitive:
    a. Information provided by the assistant should be easy to understand and interpret.
    b. Displays should provide clear, well-organized, and uncluttered information.
        i. Information should be visible in all lighting conditions, including bright sunlight and dim cabin lighting.

   c. aural cues should be easy to hear in expected cockpit environments.
   d. Haptic cues should be easily distinguished from other cues and the nominal conditions of the expected cockpit environment.
2. User-Centered design:
   a. The HMI should consider the pilot's workflow, information needs, and cognitive capabilities during the approach phase. The HMI should be easy to learn and use.
   b. The HMI should follow established industry standards and be consistent with other avionics systems to ensure clarity and usability.
   c. The HMI should offer customization settings and preferences to suit individual preferences and requirements (such as alternate formats for visual items, aural alert volume adjustments, and override modes).
   d. Elements of the HMI requiring touch should be easily reachable (such as to alert acknowledgment or overrides)
3. Real-time feedback:
   a. The HMI should provide real-time feedback to the pilot regarding the digital assistant's status and any necessary corrective actions.
   b. The HMI should provide feedback to the pilot regarding changes towards an unstable approach
   c. The HMI may also provide feedback to the pilot regarding improving approach conditions (i.e., regaining a stable approach)
4. Alerts and Warnings:
   a. The HMI should provide alerts to the pilot (e.g., warnings, cautions, & advisories, "WCAs") to signal deviations from the planned approach characteristics (e.g., glide path, airspeed, altitude, flight path)
   b. The HMI should also suggest corrective actions to regain the desired approach profile
   c. HMI alerts and messages should be clear, concise, and timely
      i. Visual, aural, and haptic alerts related to the unstable approach digital assistant should be easy to understand and distinguish from other information in the cockpit
   d. Existing cockpit system schemes should prioritize the UA digital assistant alerts and warnings
      i. Colors, volume levels, and other prioritization characteristics of alerts should be appropriate given the severity and urgency of the situation
5. Cognitive Workload:
   a. The digital assistant's HMI should not increase the overall cognitive workload of the pilot when active.
   b. The HMI design should minimize the pilot's cognitive load by presenting information logically and intuitively and reducing the number of steps required to complete tasks.
6. Integration with aircraft systems:
   a. The UA digital assistant's HMI should be integrated into existing avionics systems to enable seamless communication and coordinated actions between the UA digital assistant and the aircraft.
   b. The HMI design should ensure independence from and prevent interference with critical aircraft systems.

Funded by the European Union

    c. The UA digital assistant should ensure that higher-priority information is clearly visible and audible at all times.

7. Minimalism:
    a. The UA digital assistant HMI should aim to minimize the information given to the pilot(s) at any time
    b. The HMI design should aim to minimize any changes to the cockpit experience during an approach
8. Reliability:
    a. The digital assistant design should meet levels of robustness and reliability expected for similar cockpit assistant systems (e.g., ROPS)
    b. The digital assistant design should include fail-safes and backup routines to ensure safe and reliable operation in case of errors or malfunctions.
    c. The HMI design should include system status messages to alert the pilot to a degraded UA digital assistant situation.
9. Training and documentation:
    a. The UA digital assistant HMI should be well-documented in training materials and manuals, including descriptions of its operations, capabilities, and limitations.

AI-Based Items:

1. Explainability:
    a. The UA digital assistant should be designed such that post-flight explanations of HMI behavior (e.g., messages, tones) can be generated.
    b. Post-flight explanations of HMI behavior should be understandable to the pilots and other interested parties without knowledge of artificial intelligence (AI) or machine learning (ML) principles.
    c. HMI behavior should be predictable and in accordance with information presented in the training materials and manuals.
2. Transparency:
    a. The UA digital assistant should provide post-flight information describing the basis for in-flight alerts and recommendations (e.g., algorithms used, relevant data).
    b. The UA digital assistant should provide post-flight information describing relevant limitations and uncertainties of the models used.
3. Confidence and uncertainty estimation:
    a. The digital assistant should calculate real-time confidence levels associated with HMI behavior.
    b. Confidence levels and uncertainty associated with HMI behavior should be stored in non-volatile memory to allow for post-flight analysis.
    c. Post-flight information published by the digital assistant should identify leading factors affecting levels of confidence or uncertainty for algorithms used throughout the flight.
4. Dynamic adaptation:
    a. The digital assistant should automatically adjust visual, aural, and haptic elements to correspond to external factors (e.g., sunlight, higher vibrations, etc.).
    b. The digital assistant should be able to function correctly after overrides have been selected via the HMI.

   c. The digital assistant should comply with overrides demanded by pilots via the HMI, eliminating the use of the selected elements in alert-generating algorithms.

5. Collaborative decision-making:
   a. The HMI should support collaborative decision-making between the users and the digital assistant.
      i. The HMI should be designed according to user preferences.
      ii. The HMI should allow for in-flight user adjustments and overrides.
6. Feedback and evaluation:
   a. The digital assistant should provide pilots with post-flight feedback and evaluation of HMI behavior and their interactions with the system

Funded by
the European Union

2 Methodology

This deliverable aims to describe how ML-based algorithms for go-around predictions can be integrated into the operation. The literature review on ML algorithms in sections 1.1 and 1.2 summarized what is feasible from a technical perspective. The main task of this deliverable is to investigate, together with the relevant stakeholders, a potential concept of operation and thereby collect an initial set of high-level user requirements, which will guide the integration of the ML algorithm in the research simulator, foreseen in task 4.2. Therefore, we base the methodology for this deliverable on two pillars. Work Package 2 investigates Novel Approaches to HF, safety, and resilience in automation in this project's scope. Its D2.1 Design principles for digital assistants and HF assessment methodology (work in progress, yet to be published) provide a framework that serves as one foundation of this deliverable. Additionally, we take into account the guidance material for AI from EASA.

The regulation and certification of Artificial Intelligence in the aviation domain are currently under development and have yet to be made possible. In 2022, EASA published the EASA-AI-Roadmap, which provides guidelines for implementing Artificial Intelligence in aviation, thereby raising high-level objectives to be met for certification. The Roadmap has been complemented by two guidance documents, targeting Level1 and Level2 machine learning applications, defining Objectives and Means of Compliance for:

- Trustworthiness analysis
- Learning assurance
- Explainability
- Safety risk mitigation

The following subsections briefly summarize both documents' relevant methods and objectives for defining the Stabilized Approach Digital Assistant use case.

2.1 Human factors design principles for stable approach digital assistant

Human factors are essential when developing new tools, primarily in a safety-critical environment such as an airliner's cockpit. The tool should be visible and draw the pilot's attention when needed, but in the meantime, it should not be distracting and overwhelming. Additionally, it must be integrated into existing workflows and the working environment of present-day airliners. We apply the human factors design principle framework to achieve these objectives, which will be documented in D2.1 of SafeTeam.

Figure 8 summarizes the framework (still a work in progress) as it was available when working on this deliverable. Based on an initial idea, which for this case study was presented in the project proposal, the framework consists of three major steps:

1. System Model – Modelling the current system: This step analyses the existing system and explains why the current system should be adapted or an additional system/feature should be integrated. This step should reflect the intended or expected effects of changing the existing system. Figure 7 illustrates the information the system model shall provide.

2. Allocation Model – Designing the future system: The overall idea of this step is to define a task allocation of the adapted/extended system, which maximizes the performance of the collaborative system. Based on the system analysis of Step 1, the adapted system is modeled. The focus is on how the introduced change ripples through the components of the existing

Funded by the European Union

system. By comparing the current system's model with the adapted system's model, this step shall safeguard against potential hazards. The output of this step is a definition of what shall be implemented.

3. Implementation – Supporting enabler for human-autonomy teaming: Having understood what to change, this step defines how the new functionalities shall be implemented. Having done the technical implementation, the output of this step provides the evaluation of the implementation against the hypotheses defined in Step 1.

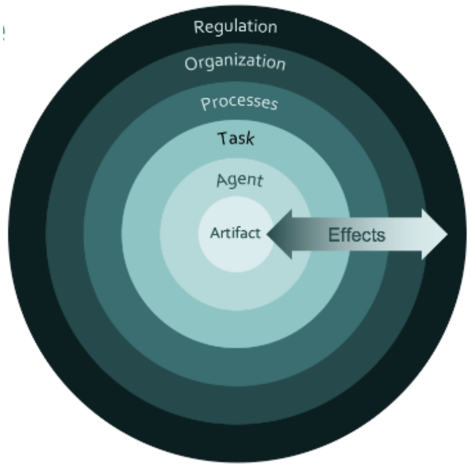


Figure 7: Information in the System Model and Allocation Model

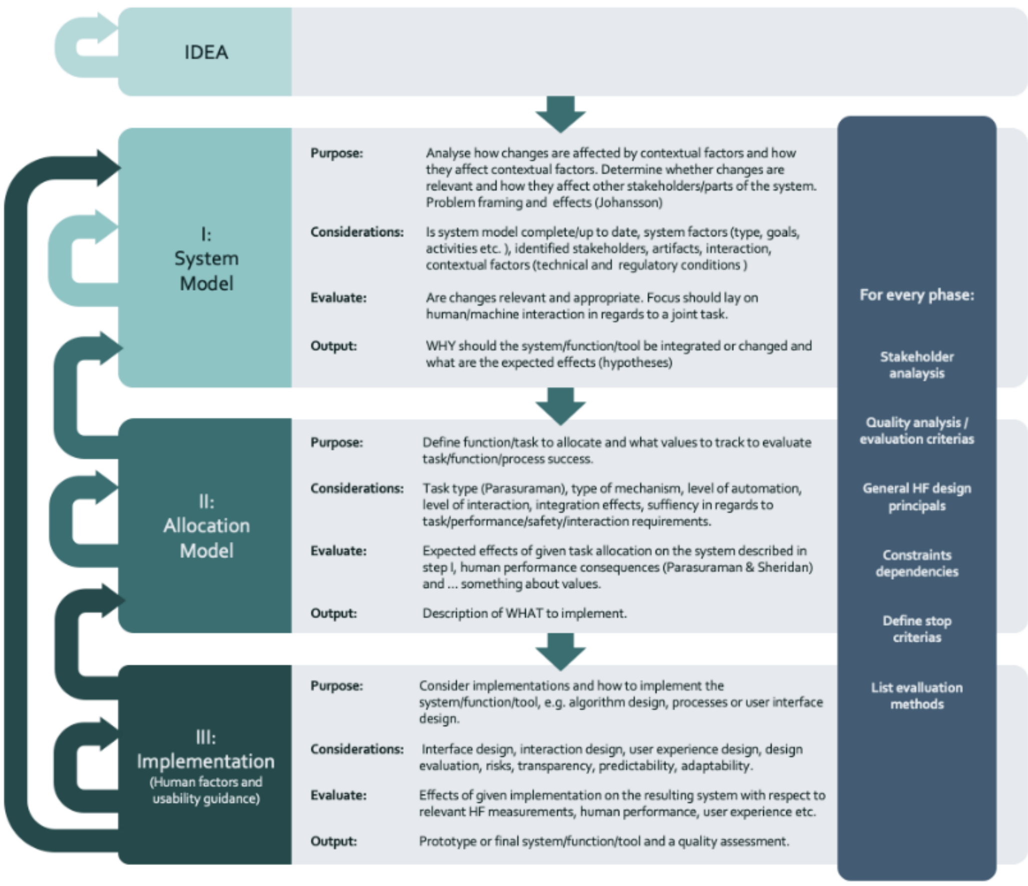


Figure 8: SafeTEAM Framework for Desing and Human Factors Assessment of Digital Assistants

Funded by the European Union

2.1.1 Methods

To accomplish the framework's steps, deliverable D2.1 proposes two methods, Hierarchical Task Analysis (HTA) and Tabular Task Analysis (TTA). By comparing the task analysis of the existing with the envisioned system, the risks of changing the system shall be understood, and ultimately, safety requirements for the new system design can be derived.

Hierarchical Task Analysis

Hierarchical Task Analysis (or HTA) is a method used to model activity. As a basis, implementing such a method requires the realization and analysis of interviews, allowing the collection of operational expertise from relevant stakeholders. An HTA describes and compares the pilots must perform to meet a predefined objective (Diaper and Stanton, 2003). The HTA method decomposes tasks into subtasks, considering the temporal relationships (sequential, parallel, or alternative) between subtasks and sometimes even the tools needed to accomplish the task and meet the objectives.

HTA aims to analyze complex systems or processes by breaking them down into smaller, more manageable components. This involves creating a hierarchical structure of tasks, where tasks are recursively broken down into subtasks up to a level in which tasks are simple enough to be easily understood.

Tabular Task Analysis

The TTA is a tabular version of the HTA. It allows tracking of the tasks identified to change with the digital assistant's introduction. Risk analysis and risk mitigation strategies are developed for each task that is either changing or introducing new. Therefore, the TTA is the initial step to introducing Safety/Risk Assessment for the Digital assistant in a structural manner.

2.1.2 Stakeholder and User Feedback

The SafeTeam framework heavily depends on the feedback of the stakeholders and users of the envisioned developments. To incorporate their domain knowledge and expectations in the case study, six workshops with pilots and other airline personnel were organized during the work on task 3.2. Depending on the number of participants, we either conducted the workshops as semi-structured interviews (if the number of participants was three or fewer, which allowed protocolling discussions) or as a mix of presentations and questionnaire sessions to enable the documentation of feedback from a larger participant group. The D2.1 provides guidance material for both activities, which was used to prepare and organize the questions, such that the workshops provide the necessary feedback from the participants regarding the activities demanded in the SafeTeam framework, steps 1 and 2. In the following, we summarize the objectives and high-level topics of the two activities. A complete set of questions and the minutes of the workshops will be provided in Appendix.

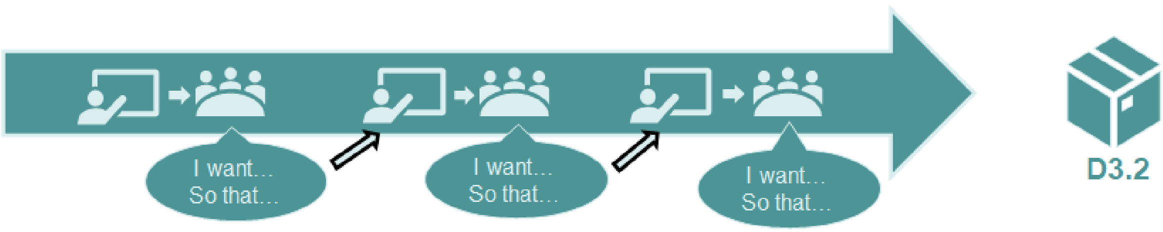


Figure 9: SafeTEAM Workshops for Stakeholder and User Feedback

2.1.3 Semi-Structured Interviews

The organization of the workshop included three phases:

- Preparation is explained in more detail in Table 10
- Execution
- Post-Processing, explained in more detail in Table 11

After each workshop, we analyzed the results and included them in the presentations for the following workshops. We, however, did not present the findings at the beginning of the upcoming workshops but started with a general discussion to allow for fresh ideas. The output of the workshops is a set of user stories, which are the basis for the objectives, system boundaries, high-level functionalities, and HMI findings, presented in section 3.3.1.

Table 10: Workshop Preparation

<table>
<tr><td rowspan="4">Preparation</td><td>Setting:</td><td>Three workshops were performed in person, three workshops were performed online via Zoom.
Four workshops (workshops 1, 2, 3, and 6) were scheduled for 2 hours, one workshop (4) was scheduled for one day and one workshop (5) was scheduled for 1,5 days.</td></tr>
<tr><td>Goal:</td><td>The goal of the workshop was to get feedback from end-users and stakeholders, on how to integrate machine-learning-based unstable approach predictions, as developed in SafeClouds.eu and described in section 1.2, into the cockpit. Focus was put on:
• expected benefits
• expected negative effects
• existing systems preventing unstable approaches
• boundaries of the envisioned digital assistant</td></tr>
<tr><td>Participants:</td><td>The interviewer side participated with two persons. This allowed consistent documentation while keeping the workshop in flow.
For the semi-structured interview, we organized workshops with one - three participants.</td></tr>
<tr><td>Discussion Topics:</td><td>In each workshop, we structured the Question into the categories:
• General
• Operation
• HMI
• Machine Learning

Before each set of questions, a presentation introducing the workshop participants</td></tr>
</table>

Table 11: Workshop Post-Processing

<table>
<tr><td rowspan="2">Post Processing</td><td>De-Briefing</td><td>After the workshops, the team discussed the interviews and checked the meeting minutes which were written during the workshop.</td></tr>
<tr><td>User Stories</td><td>The feedback/answers of the participants were reformulated in user stories and grouped into the categories:
• Objectives
• High-level functionalities / Operation
• HMI
• Machine Learning</td></tr>
</table>

Funded by the European Union

		In Appendix B Stakeholder and User Workshops/Interviews, each Workshop is documented and the user stories from each workshop are presented

2.2 Regulatory Guidance for AI applications in aviation

EASA's Concept Paper: First usable guidance for Level 1&2 machine learning applications [16] defines a framework for building trustworthy AI. It contains a set of objectives grouped into four categories, as illustrated in Figure 10. While the complete document is relevant for this project (especially when implementing and evaluating the case study in Work Package 4), the case study definition phase described in this document, especially the "Characterisation of AI" objectives, are of importance

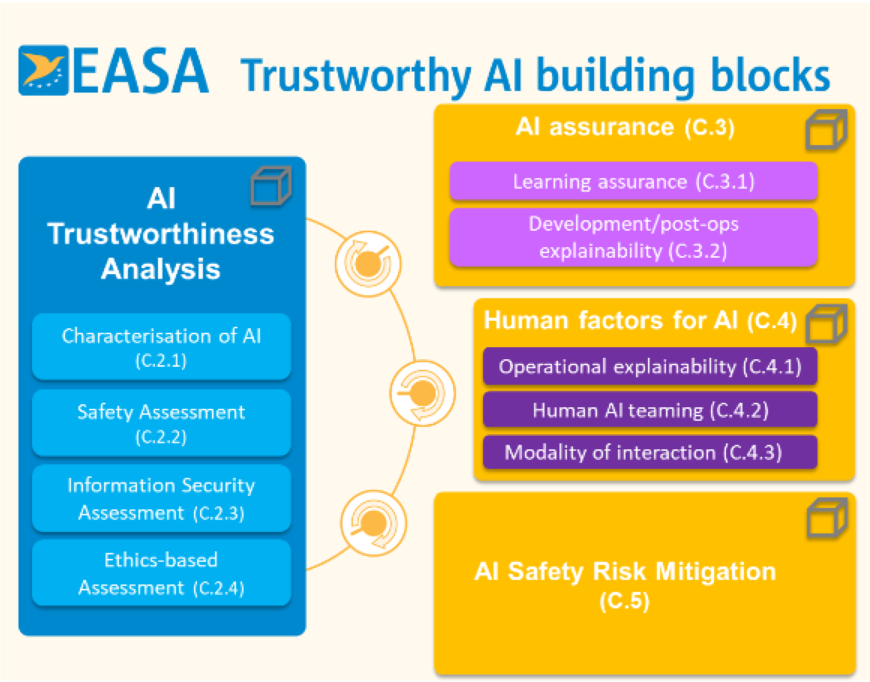


Figure 10: EASA's Trustworthy AI Building Blocks [16]

.

2.2.1 Characterization of the AI application

For the Characterization of the AI, EASA defines the following objectives, summarized in Table 12. Comparing EASA's goals with the framework proposed in D2.1 (specifically Step 2: allocation model), we find several thematical similarities in both approaches. Objectives, by definition, must be complied with. EASA, as a regulatory Agency, leaves the means of compliance to the developer. Thus, applying the proposed framework of D2.1 for defining this case study and evaluating if the results satisfy the EASA's objectives is an interesting side effect of task 3.2. Objective CO-04 is the most challenging and open objective. Producing a ConOPS is not foreseen per se in the SafeTEAM framework. However, a substantial overlap can be found when comparing the objectives of the SafeTEAM framework, especially the System Model and Allocation Model, with the EASA Guidance and, e.g., the outline of the US Justice Department for compiling a ConOPS. In our view, a detailed description of the System Model and Allocation Model, combined with a set of high-level requirements, satisfies the US Justice Department's guidance for a ConOPS, which can be found in Appendix A.

Table 12: EASA's Objectives related to ConOPS [16]

EASA Objective ID	Objective
Objective CO-01	The applicant should identify the list of end users that are intended to interact with the AI-based system, together with their roles, their responsibilities, and their expected expertise (including assumptions made on the level of training, qualification, and skills).
Objective CO-02	For each end user, the applicant should identify which high-level tasks are intended to be performed in interaction with the AI-based system.
Objective CO-03	The applicant should determine the AI-based system taking into account domain-specific definitions of 'system'.
Objective CO-04	The applicant should define and document the ConOps for the AI-based system, including the task allocation pattern between the end user(s) and the AI-based system. A focus should be put on the definition of the OD and on the capture of specific operational limitations and assumptions.
Objective CO-05	The applicant should document how end users' inputs are collected and accounted for in the development of the AI-based system.

Funded by
the European Union

3 Stabilized Approach Digital Assistant – Use Case Definition

3.1 System Model

This section describes the system model for the stabilized approach digital assistant case study. The system model describes the current approach phase of aircraft **without** the envisioned machine-learning-based unstable approach prediction artifact/constituent. The system model contains information on the system:

- boundaries
- stakeholders
- processes
- tasks
- agents
- artifacts

Figure 11 illustrates a schematic overview of the stabilized approach digital assistant's system model. The purpose of the system model is to provide a systematic description of what must be considered when enhancing the existing system with a new functionality/artifact (in our case, ML-based unstable approach prediction). For the stable approach digital assistant case study, an aircraft's approach phase is considered a system subject to change. In the following subsections, each level of the system model is described in more detail.

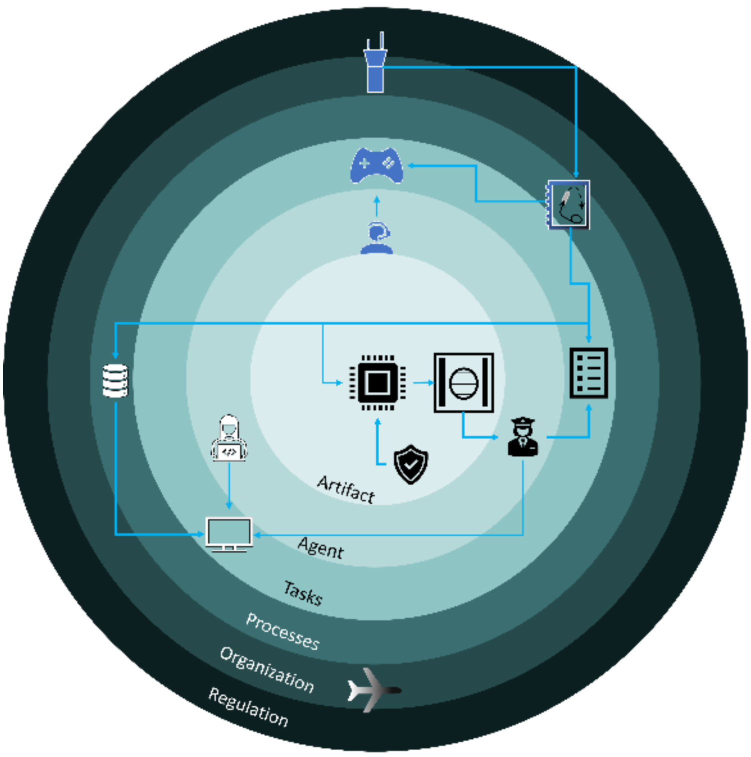


Figure 11: System Model Components

The approach phase has several **goals**. Primarily, the approach phase shall prepare a **safe landing**. Unstable approaches are a precursor for increased risk during a landing and are a quantifiable Key Safety Indicator for airlines. Thus, **avoiding unstable approaches or maximizing stable approaches** is considered a **primary goal** of airlines. In case of unstable approaches, go-arounds shall be executed by the cockpit crew. Maximizing **the ratio of go-arounds in case of unstable approaches** is also **a safety target** for airlines. **Optimizing time and fuel consumption** during the approach phase are **secondary goals** that are, however, directly linked to the reduction of unstable approaches and, consequently, go-arounds.

3.1.1 System Boundaries

In this section, we define the (operational) system boundaries for the stabilized approach assistant in terms of a generalized workflow of the cockpit crew during approaches based on user stories NF.5 and NF.3. Figure 13 provides a summary of milestones during the approach. The management of the energy state of an aircraft starts at the descent point. The distance needed to dissipate energy largely depends on the cruise altitude. Usually, the descent is managed by the flight management computer to be as efficient as possible. Even though the chosen descent point can decide if an approach becomes unstable, it usually is during the final approach when the success of an approach is determined. Depending on the airport approach procedures, the final approach begins between 2000 ft AGL and 4000 ft AGL followed by the Gear Extension and the Final Flap Setting.

Table 13: List of Approach Milestones, highlighting which are within the System Boundaries

Approach Significant Points	Altitude Range	Track Mileage	Speed Range
Start Descent	Cruise Level	FMS Descent Path (~100-130 NM)	Cruise Speed
Start of Flap Config.	3500 - 4500 ft AAL	12 - 15 NM	200-220 kts IAS
Intercept Glide Slope	2000-4000 ft AAL	6 - 11 NM	180 kts IAS
Gear Extension	2000 ft AAL	6 NM	160 kts IAS
Final Flap Setting	1200 - 1500 ft AAL	4 NM	approach speed
Stabilized	1000 ft AAL	3 NM	approach speed

Figure 12 illustrates the approach phase, beginning ~10NM from the runway threshold. The aircraft intercepts the glide slope from below. Furthermore, the figure shows the significant points of the approach phase, the final approach fix, and the 1000ft gate. Figure 12 and Table 13 are only generalized descriptions of an approach. Each runway has procedures that specify the approach. An example procedure is provided in the process section below.

Funded by
the European Union

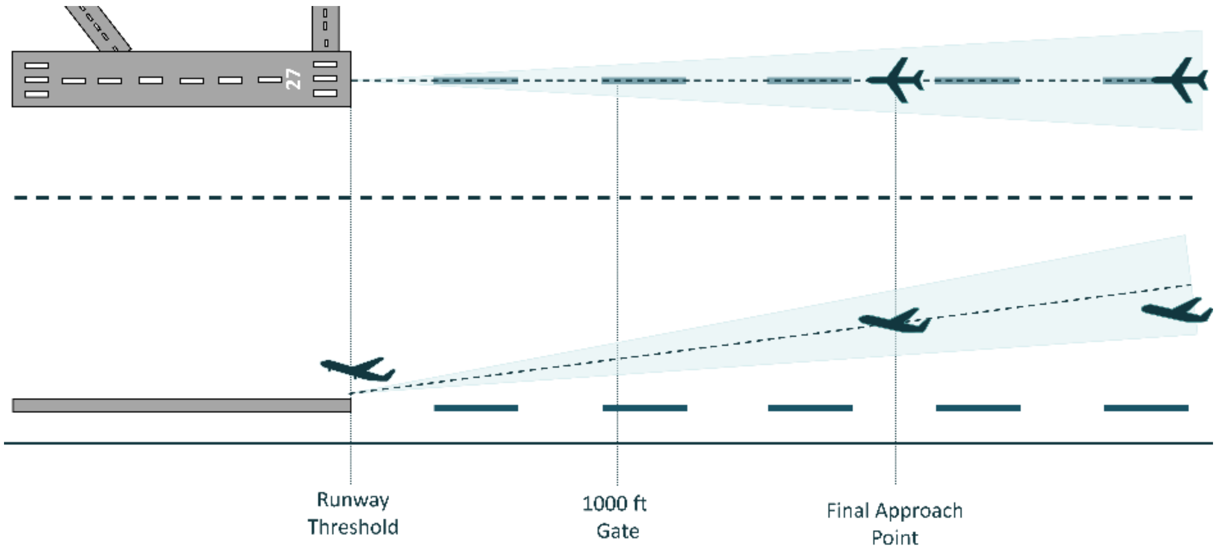


Figure 12: Approach Phase: System Boundaries

3.1.2 Stakeholders / Agents

In workshops and from the consortium composition, we were in close contact with two relevant stakeholders:

- Airline/Operator: is the most relevant stakeholder in this system model. They operate the aircraft fleet and define applicable procedures like stabilized approach criteria. Furthermore, they determine the pilot tasks in the approach phase to an extent, besides the aircraft manufacturers, via training. Additionally, the Flight Data Monitoring departments are involved in evaluating the safety of an approach. From the airline, the relevant agents are:
    - Pilots
    - FDM Experts
- Regulatory Bodies: are involved in the procedure designs of airlines and air navigation service providers.

Additional relevant stakeholders are:

- Air navigation service providers (ANSPs): They operate/control the airspace according to procedures (relevant ones introduced below) agreed upon with the regulatory bodies. From the ANSP, the appropriate agents in the system are:
    - Air Traffic Controllers
        - Tower Controller
        - Approach Controller

3.1.3 Processes / Procedures

Various options exist for conducting an approach, depending on the available infrastructure and navigation aids at an airport. Besides an ILS Approach with vertical guidance (3D approach/precision approach) installed at most major airports and runways, there are also approaches without vertical guidance (2D approach/non-precision approach) based on GPS, Localizer only, VOR, or NDB. These 2D approaches typically increase the workload for the cockpit crew. This could be seen as a further influencing factor on unstable approaches.

3D Approaches

EASA defines a three-dimensional approach as "... an instrument approach operation using both lateral and vertical navigation guidance." Unlike the more commonly known precision approach, three-dimensional includes typical precision approaches such as ILS and LPV. Still, they also include approaches that only provide lateral guidance, but the flight management system provides vertical guidance. For example, take a VOR/DME approach, where the ground navigation aid only provides lateral guidance. However, the flight management system of modern aircraft can calculate appropriate vertical guidance.

2D Approaches

Two-dimensional approaches, on the other hand, are defined as "... instrument approach operation using lateral navigation guidance only". Therefore, the pilots perform the vertical guidance manually by crosschecking distance and altitude from approach charts. This significantly increases workload compared to a 3D approach.

Missed Approach Procedure / Stabilized Approach Criteria

A relevant procedure, which shall be considered in the allocation model, is the missed approach procedure. The standard missed approach procedure is defined within the approach procedure chart. One reason to initiate a missed approach procedure is an unstable approach. The exact definition of an unstable approach depends on the airline. Every airline has to define stabilized approach criteria in its Standard Operating Procedures (SOPs), and unstable approaches are those approaches that do not meet these criteria. Some guidelines exist on how to classify unstable approaches, e.g., from the Flight Safety Foundation (p44.) [7] or EASA's Data4Safety Project. However, as long as there are diverging SOPs, from our experience, one has to apply the airline-specific rules for classifying unstable approaches since these are what the pilots are trying to fulfill. Using general guidelines, especially when labeling data for a machine learning application, will result in either too high or too low unstable approach rates from an airline perspective.

3.1.4 Task Analysis

For the pilots as end-users in the identified system, Figure 13 illustrates their tasks in the boundaries defined above. The task analysis results from the workshops performed with pilots for task 3.2. It shall provide a general outline of the tasks that have to be fulfilled by both pilots throughout the final approach, independent of aircraft type or operator-specific procedures.

Funded by
the European Union

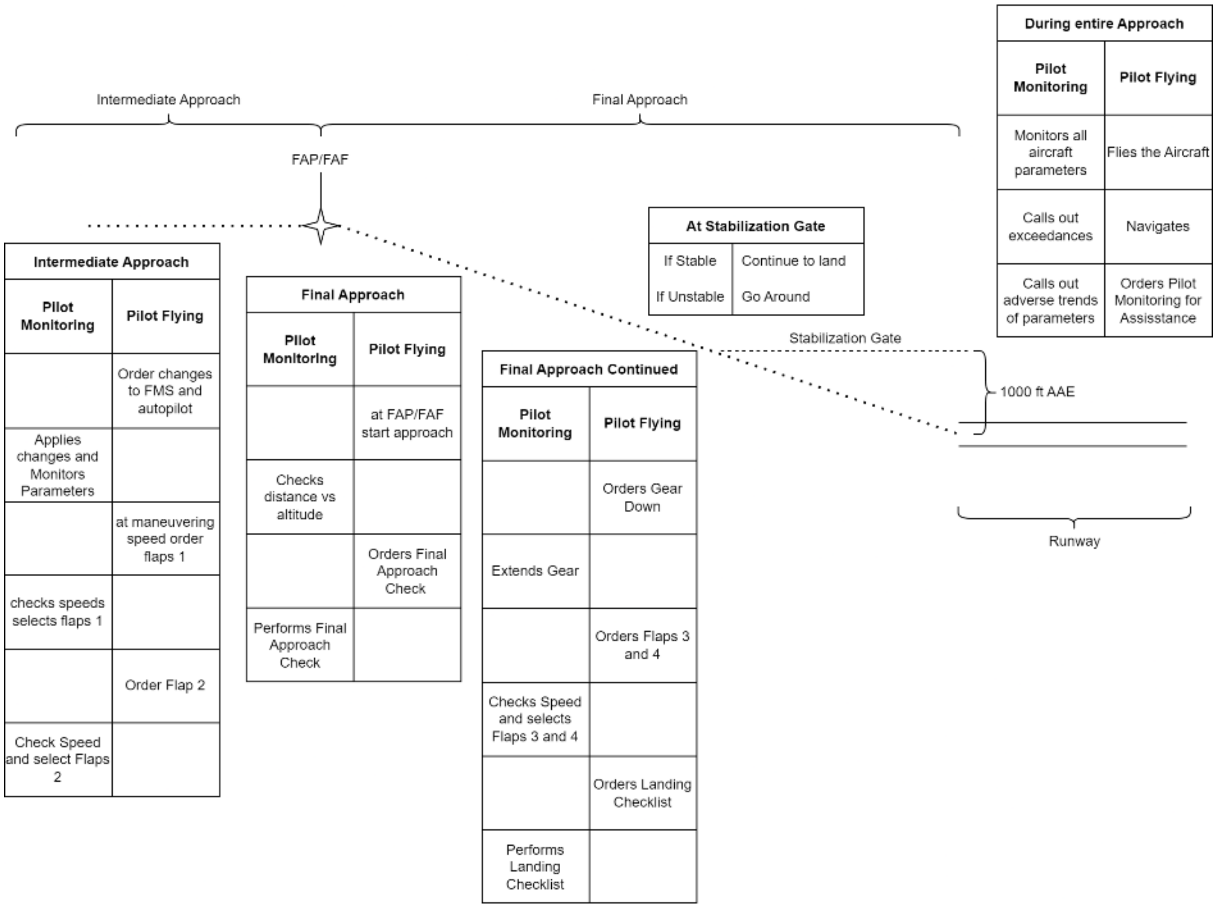


Figure 13: Sequential Task Analysis of the Approach Phase

Additionally, Figure 14 provides a Hierarchical Task Analysis of the approach phase system. This diagram groups the tasks according to the objective they aim to achieve. The hierarchical structure becomes ever more abstract from the bottom to the top.

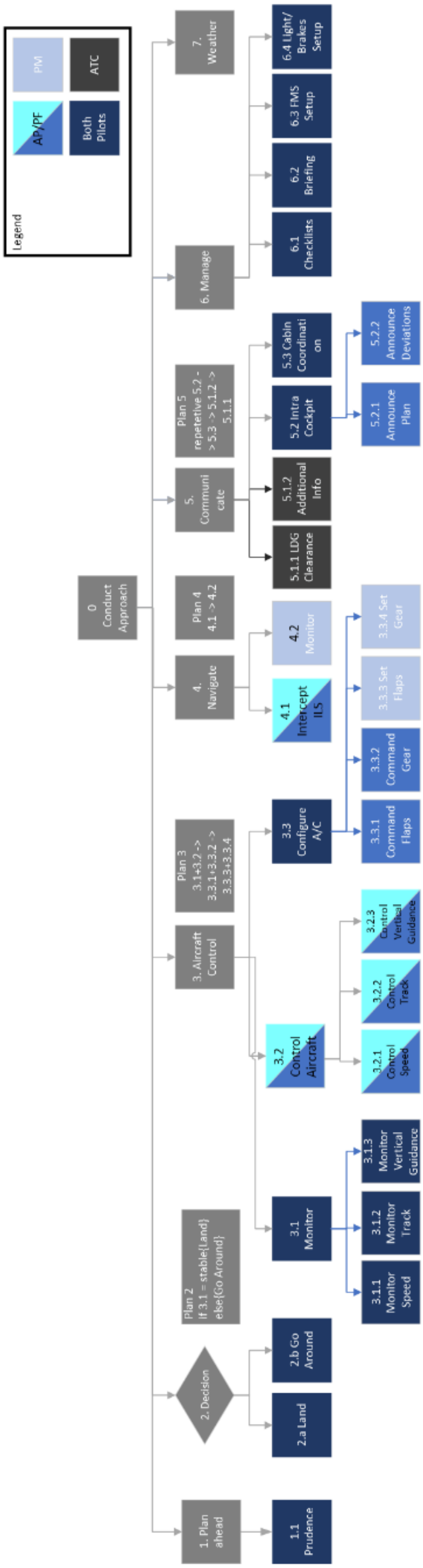


Figure 14: Hierarchical Task Analysis of the Approach Phase

Funded by the European Union

3.1.5 Artifacts

The cockpit Human-Machine Interface (HMI) is the relevant artifact in this case study. The following sections discuss the current state of the art – first in terms of typical HMI elements present on commercial airliners and then in terms of existing digital assistants and related systems (e.g., ROPS, TAWS). While systems to prevent runway-related incidents already exist, there are notable differences between current capabilities and desired safety outcomes. The existing designs, their HMI functionality, and gaps in capabilities are described below.

Human-Machine Interface (HMI)

The HMI in this context comprises a combination of visual, aural, and haptic cues, as well as relevant links to the cockpit systems that would provide those cues to the flight crew. The diagram in Figure 15 illustrates possibilities for HMI elements on a commercial airliner (similar to a Boeing 737 or Airbus A320) [17] [18]. While each airline and aircraft model varies in cockpit hardware, software, and design philosophies, common elements are plentiful and more than sufficient for designing a stabilized approach digital assistant.

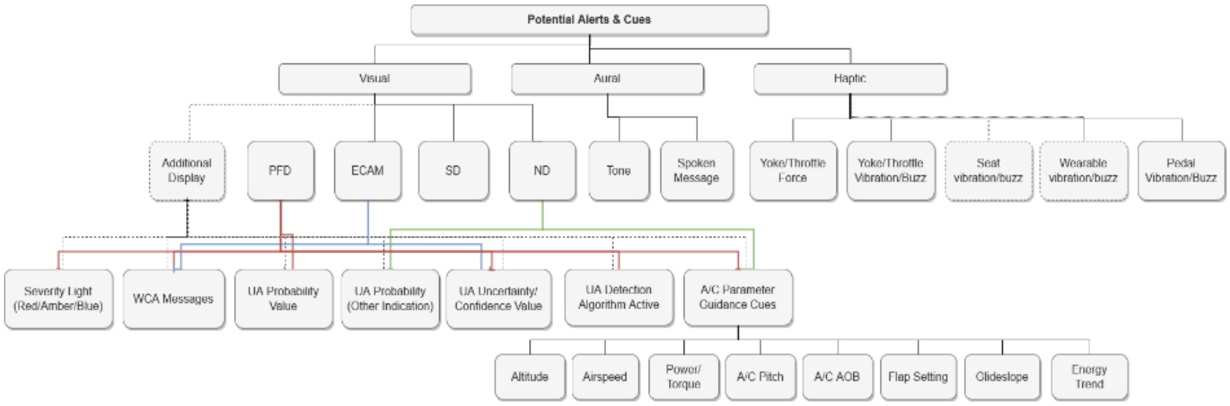


Figure 15: Stabilized Approach Digital Assistant HMI Options

The goal of this diagram is to capture all the options possible for communicating alerts and information generated by the digital assistant. These options are organized based on sensory channels, the path of delivery, and then the type of information/indication/alert presented to the crew - followed by further detail, as needed (e.g., parameters of interest). In this diagram, the more complex or unconventional options were indicated with a dashed outline – to indicate that the research group considered these options but did not expect them to be popular among airlines and pilots. In this way, the diagram could be used to lead discussions about the digital assistant HMI design space without unpopular options distracting too much attention.

During the approach phase of flight, the Primary Flight Display (PFD) and Navigation Display (ND) provide the most critical information, which is also the most relevant for maintaining stabilized approaches. While pilots follow a method of continuously scanning various instruments and the outside environment during flight, typically, the PFD serves as the main focal point for the pilot in command (PIC) during an approach. The copilot’s attention is usually divided between the ND, communication tasks, and configuration actions. Other relevant avionics components include the Electronic Centralized Aircraft Monitor (ECAM), System Display (SD), and warning panel. An image of the overall cockpit layout for a typical Airbus A320 aircraft is shown in Figure 16, and enlarged images of other key avionics components are shown in Figure 17. [19]

Funded by the European Union

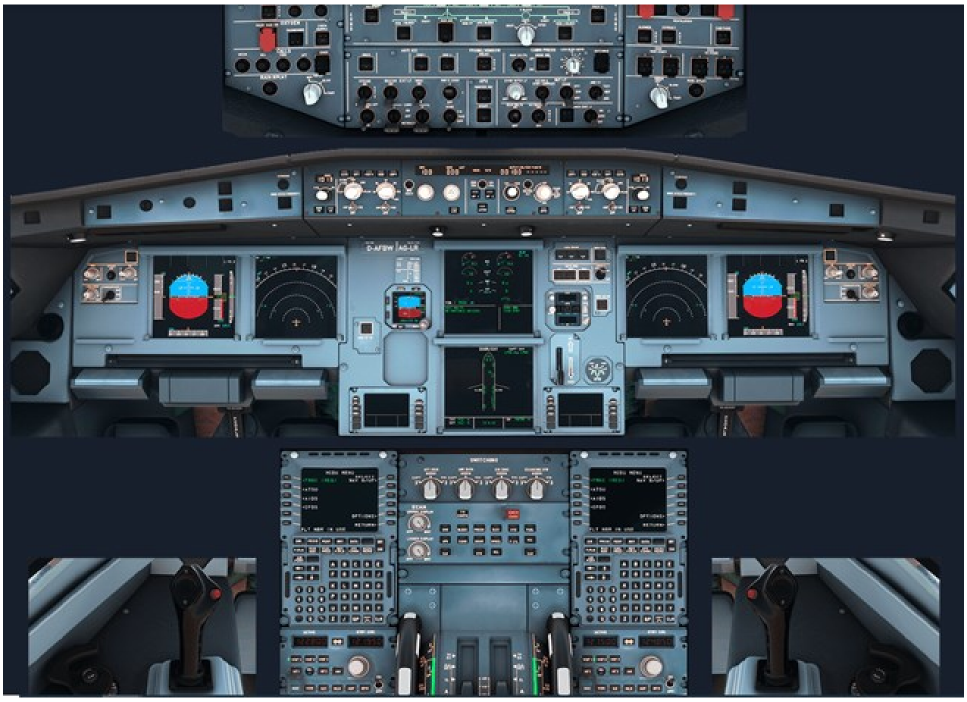
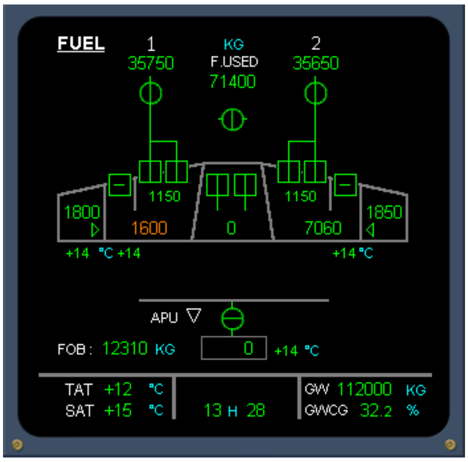


Figure 16: Example Airbus A320 Cockpit Layout. [18]

Primary Flight Display (PFD)

Navigation Display (ND)

Electronic Centralized Aircraft Monitor (ECAM)



System Display (SD)

Figure 17. Relevant A320 Cockpit Avionics Components, Enlarged. [18]

Funded by the European Union

Figure 18 shows the changing levels of workload typically seen throughout a flight. Pilots are very busy during the approach phase of flight: they must perform a constant stream of actions to prepare for landing while simultaneously interpreting a large amount of quickly-changing information—including aircraft state parameters, directives from air traffic control, other air traffic, potential ground hazards, and weather. Any additional information or aids considered for this environment must be carefully balanced against the potential for cognitive overload, confusion, and distraction. Designs considered for the stabilized approach digital assistant must be carefully evaluated to determine their effects on factors such as pilot workload and situational awareness. [13] [19]

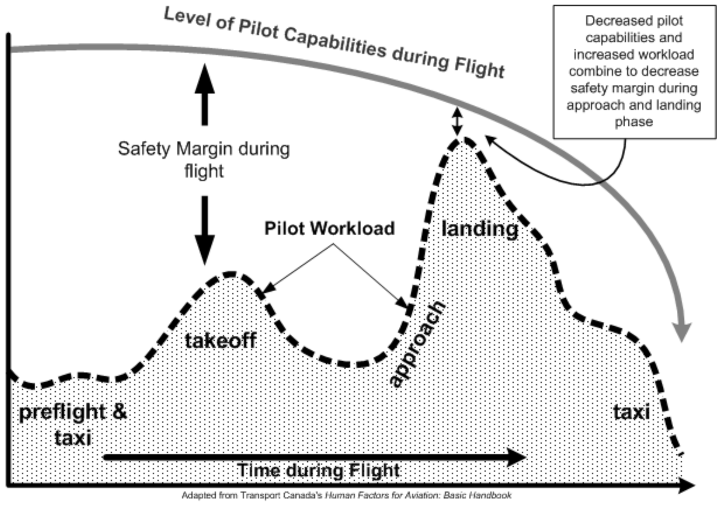


Figure 18. Pilot Workload versus Phases of Flight [19]

Runway End Overrun Warning (ROW) / Runway End Overrun Protection System (ROP)

Runway excursions result from various factors during the approach and landing phase, such as an unstable approach due to too high energy and a slippery or contaminated runway. The runway end overrun warning/runway end overrun protection system (ROW/ROP) is a progression from the Brake-to-Vacate (BTV) feature Airbus first developed for the Airbus A380 and was type certified by EASA in the year 2009 [4]. It aims to reduce the risk of a runway overrun by warning the crew during the final approach and during the roll-out of an imminent runway overrun risk. It is available on the Airbus A320, A330, A350, and A380 aircraft families. Its alerting function effectively guides and assists the flight crew in the go-around decision-making process during each approach. Additionally, it insists on applying all available deceleration means during rollout on the runway to avoid an imminent overrun. [20].

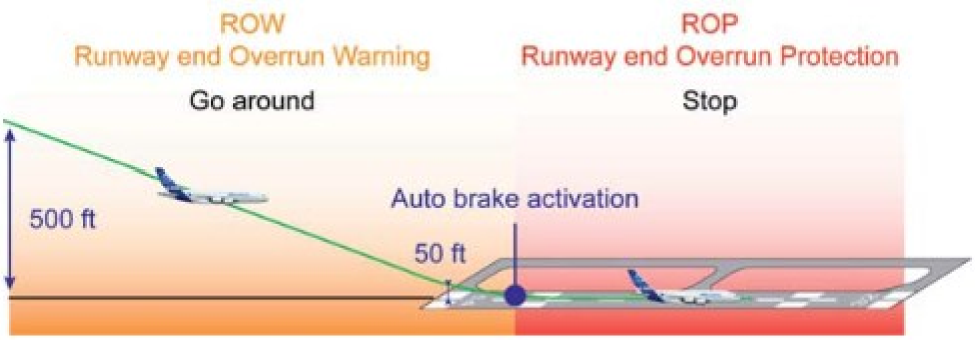


Figure 19: ROW/ROP illustration from [21]

The ROW/ROP algorithm computes eight times per second the stopping distance based on the "aircraft ́s weight, ground speed, wind condition, landing configuration, and vertical/horizontal trajectory concerning the runway threshold." It computes the actual landing distance for dry and wet runways and presents the warnings aurally and visually in the primary flight display (PFD). To identify the runway intended for landing, the system uses data taken from the terrain awareness and warning system (TAWS) database [21] [20].

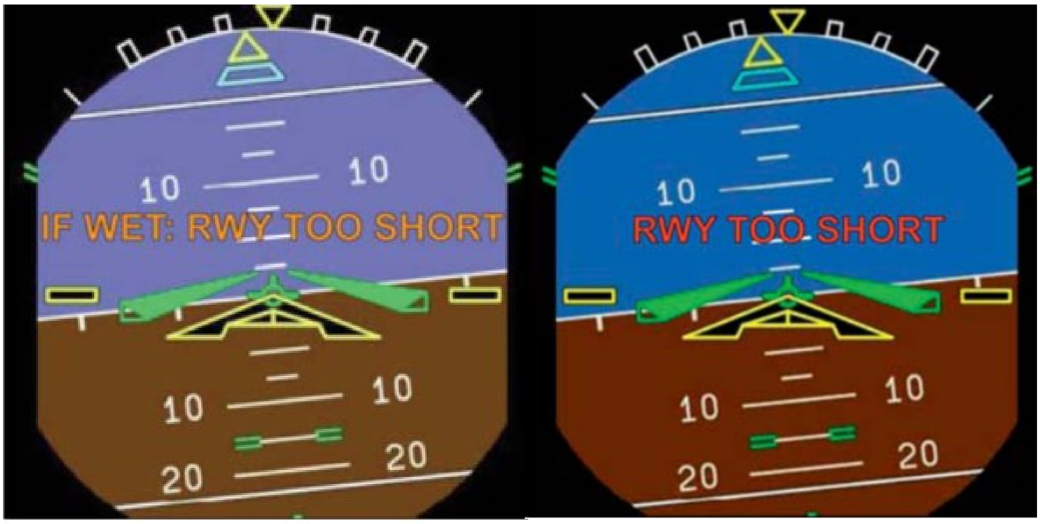


Figure 20 ROW/ROP warning in an Artificial Horizon [21]

SmartRunway and SmartLanding – Unstable Approach Monitor

Honeywell produces an extension to the Extended Ground Proximity Warning System (EGPWS), the SmartRunway, and SmartLanding. Especially the Unstable Approach Monitor is relevant for this case study. After passing the stabilization gate, the system monitors the final approach and acts similarly to a Pilot Monitoring. The system produces an aural annunciation if the aircraft is:

- too high
- too fast
- wrongly configured

between 950ft and 450ft above runway elevation and an acoustic 'Unstable - Unstable' warning if the conditions still apply below 450ft above runway elevation. Additionally, the caution and warning are also displayed in the navigation display. Figure 21 illustrates the functionality of an excessive approach angle. For more details, we refer to Skybrary.

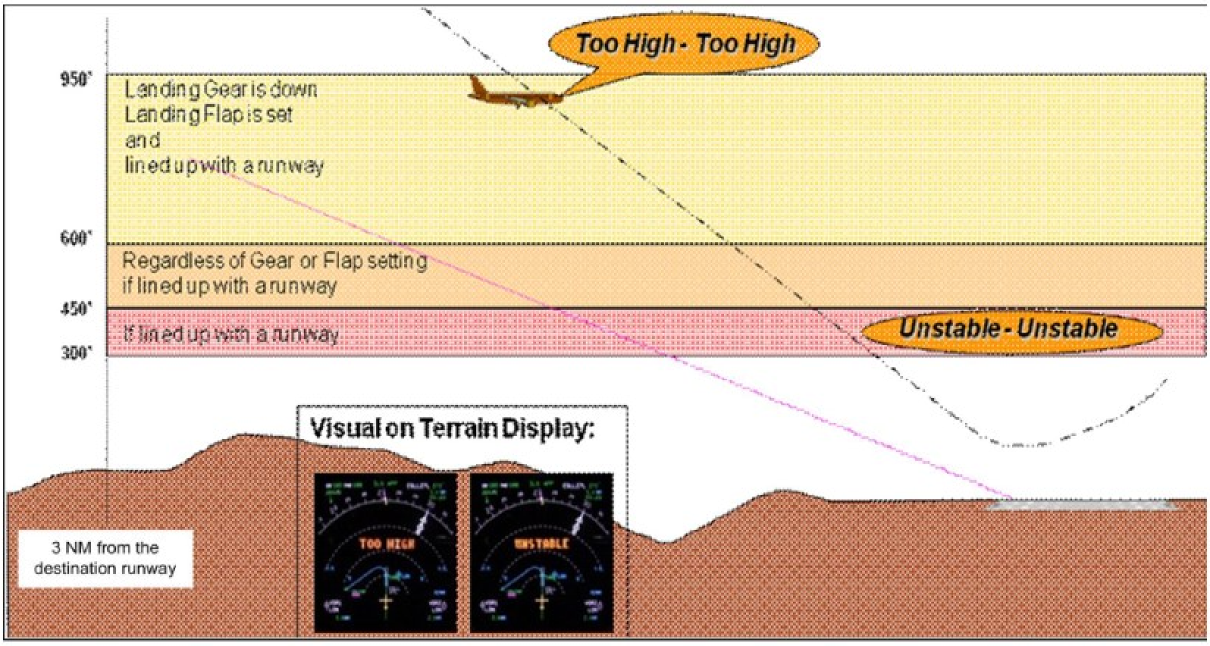


Figure 21: Illustration of Smart Runway Smart Landing Too High functionality [22]

The inhibit functionality provided by the EGPWS system is essential for later discussions. Figure 22 illustrates the GPWS panel in an aircraft. It allows setting the functionality inactive as a whole or by operating modes, in case the pilot wants to.

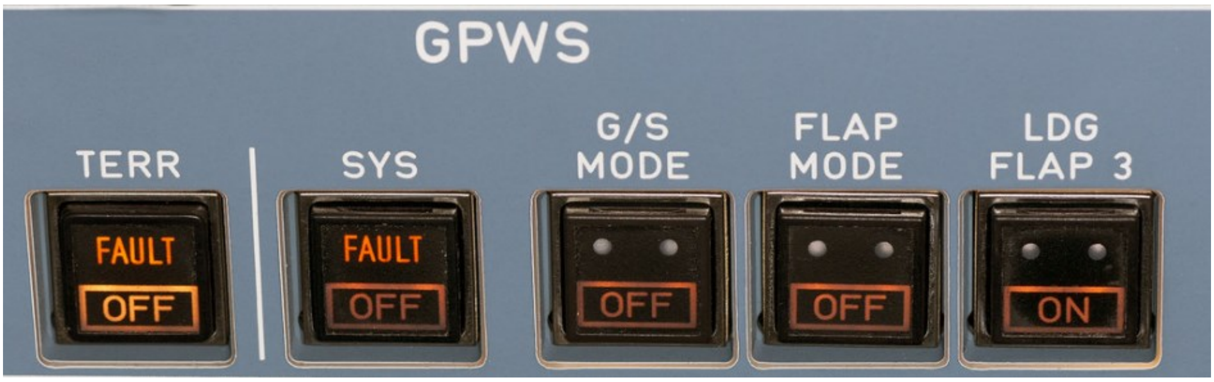


Figure 22: GPWS Panel, with Options to inhibit (partial) functionalities.

3.2 Allocation Model

While the system model describes the state-of-the-art approach phase, the allocation model in this section works out the changes that ripple through the system model by introducing an additional component, the stabilized approach digital assistant, in the cockpit. The section will first provide a high-level description of the idea behind the stabilized approach digital assistant and thereafter work through the parts of the system model affected by the envisioned modifications. This deliverable focuses on the airline-specific aspects of the system model, especially those with a cockpit relation. This is in line with the overall scope of SafeTEAM, focusing on the teaming of digital assistants and human users. As far as the technology is developed at this stage, the technological aspects have been discussed in detail already in SafeClouds.eu. We will still modify the existing technology according to the user's needs. On the contrary, one purpose of this exercise is exactly to discover user expectations regarding, e.g., potential data sources or functionalities that are not yet implemented in the machine learning constituent developed in SafeClouds.eu, and collect a set of high-level requirements that will guide the implementation phase in task 4.2.

3.2.1 High-Level Description

We aim to investigate a non-deterministic, ML-based assistance tool that provides timely information about the stability of the approach to pilots during the final approach phase. The basis for this assistance tool is the predictive functionality of the machine learning algorithms, described in section 1. Review of Machine Learning Models from SafeClouds.eu. This section defines how the machine learning artifacts could best be implemented in the approach phase, as defined in the system model in section 1. System Model. Figure 23illustrates the additional ML artifact within the schematic illustration of the system model. At this research stage, the ML component shall be used as a decision support tool, as requested in the stakeholder workshops and documented in user story HL.P.1, providing the pilots with real-time information on unstabilized approach risks during the final approach phase. Therefore, the Stabilized Approach Digital Assistant groups into EASA's AI level 1B, according to Table 14.

Table 14: EASA's Levels of Automation according to [16]

AI level	**Funciton** allocated to the system to contribute to the high-level task	Authority of the end user
Level 1A Human augmentation	Automation support to information acquisition	Full
	Automation support to information analysis	Full
Level 1B Human assistance	Automation support to decision-making	Full
Level 2A Human-AI cooperation	Overseen and overridable-automatic decision	Full
	Overseen and overridable automatic action implementation	Full
Level 2B Human-AI collaboration	Overseen and overridable-automatic decision	Partial
	Overseen and overridable automatic action implementation	Partial
Level 3A Supervised advanced automation	Supervised automatic decision	Upon alerting
	Supervised automatic action implementation	Upon alerting
Level 3B Autonomous AI	Non-supervised automatic decision	Not applicable
	Non-supervised automatic action implementation	Not applicable

The parts of the system model directly affected by the Digital Assistant are encircled inFigure 23 by the orange ellipse.

Funded by the European Union

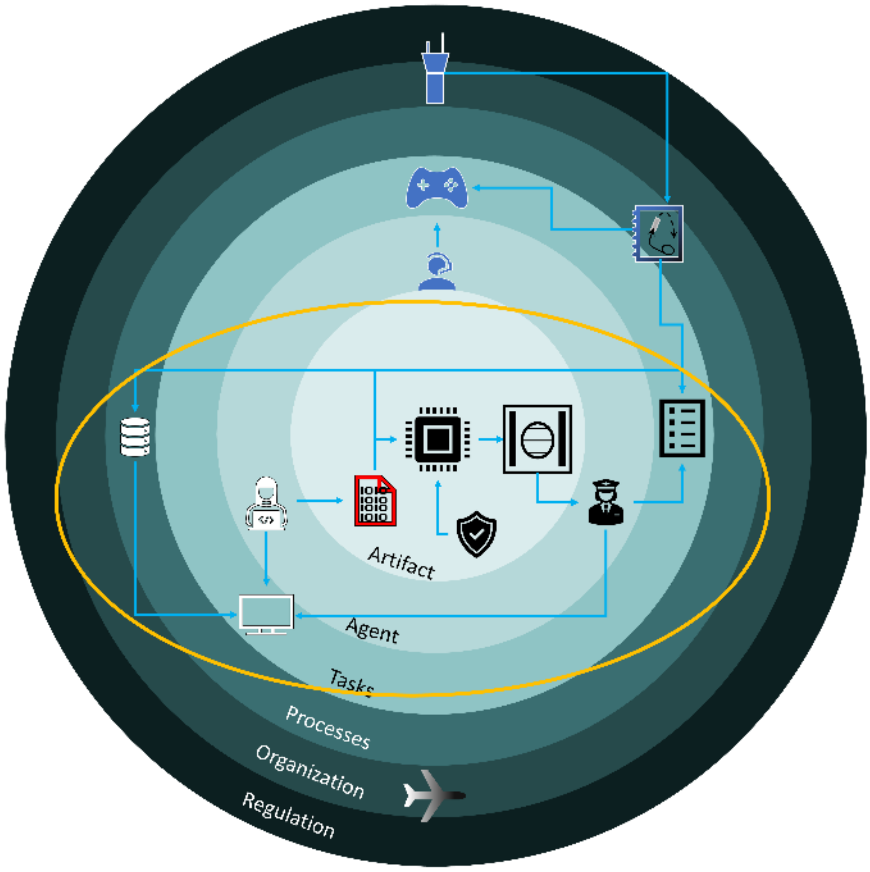


Figure 23: Schematic illustration of the Allocation Model and which components are affected by the Digital Assistant

From the airline perspective, the stabilized approach to a digital assistant can be subdivided into two parts, as illustrated in Figure 24:

- On-Board or Online part, which is the actual implementation in the cockpit and
- Off-Board or Offline part combines gathering data and training ML algorithms based on the data available to the Flight Data Monitoring (FDM) department.

The offline part is summarized in section 1. Review of Machine Learning Models from SafeClouds.eu. The complete description of an IT infrastructure, setting up a data pipeline, model benchmarking, and training is described in cite D42 / D43 from SafeClouds.eu. This deliverable focuses on the online part, where the output Off-Board process must be integrated into the cockpit. To capture changing operations and train the model on rare events, not in previous data sets, the model needs to be prepared with new data at regular intervals. Therefore the offline part of the system requires access to the airlines' data lake. The following graphic shows the information flow and provides an overview of all relevant interfaces.

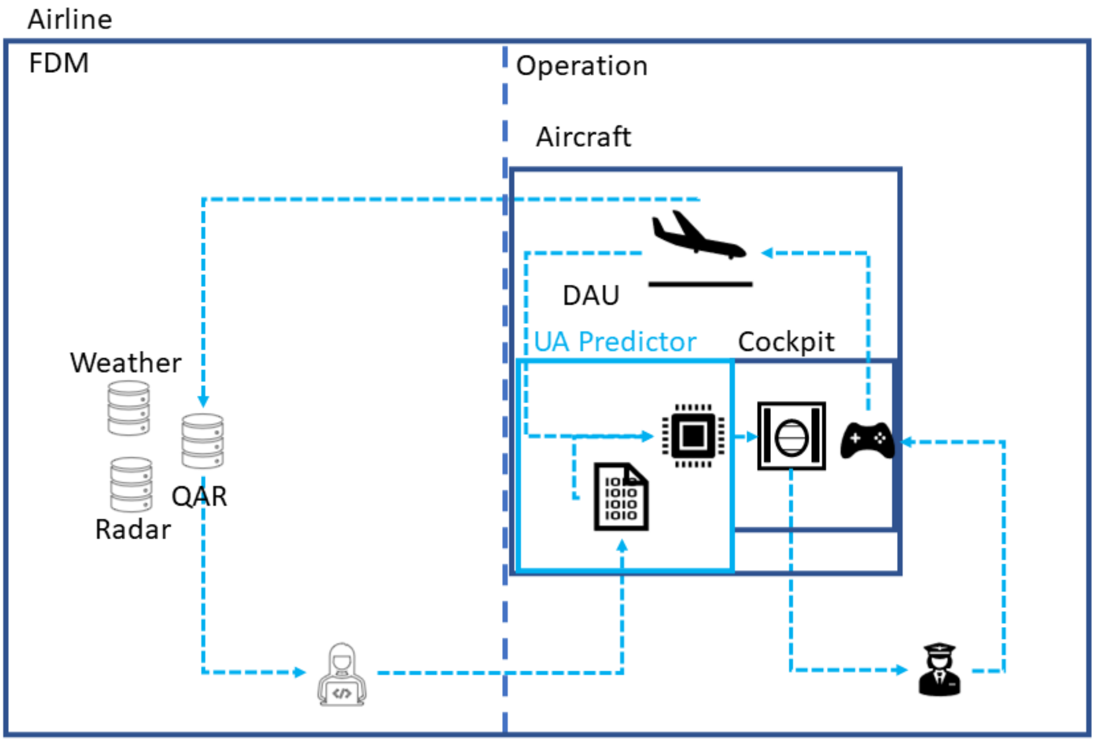


Figure 24: Online and Offline part of the Digital Assistant

Goals and Objectives

Tools exist that detect unstable approaches and consequently request pilots to perform a go-around, as described in 1. System Model - Artifacts. However, the workshops performed within this project and the flight safety foundation's investigation on unstable approaches indicate that the ratio of go-arounds in unstable approaches still needs to be higher. The idea of the stabilized approach digital assistant is to provide relevant yet predictive information regarding the stability of an approach to the pilot already before the stabilization gate. The primary purpose of the introduction of the ML-based unstable approach prediction algorithm in the aircraft cockpit is to provide a decision support tool that increases the pilot's situational awareness and hence decreases the number of unstable approaches (O.1). A further objective of a timely indication that an approach is (about to become) unstable, is the pilots' preparedness for a go-around. We expect the assistant to increase the number of go-arounds if an approach is not stabilized at the stabilization gate (usually at 1000 ft above aerodrome elevation) (O.3). A secondary objective arising when reducing unstable approaches is fuel and time savings (O.2). Additionally, ANSPs could benefit from a reduced number of missed approaches since they avoid a second approach which would reduce their capacity.

3.2.2 Task Analysis - Work Processes

From the workshops, we found that the end users desire several modes of operation based on separate functionalities. An overview of all discussed modes is presented in Table 15. A short description of the requested functionality is summarized in the user stories section 3.3.1.

Table 15: Modes of the Digital Assistant

Mode	User Story ID	Output	Starting at

Funded by the European Union

Inactive/Passive (when outside the operational domain)		none	upon Pilot input
Monitoring	HL.M.1	detection & reason for detection	Stabilization Gate
Prediction (when in the operational domain)	HL.P.1, HL.P.2	prediction & contributing factors	FAF/FAP

From the FAF/FAP until the runway threshold, a go-around risk prediction mode shall be available and active when the approach is inside the operational domain of the ML constituent. Otherwise, the **Prediction Mode** shall be inactive. The **Data gathering** necessary to compute features, which are the input to the ML-constituent, starts 15-20NM from the runway threshold, defined as the start of descent in Table 13.

The end users' tasks in the state-of-the-art approach phase are described in section Task AnalysisTask Analysis. This section analyzes how tasks could be distributed with the stabilized approach digital assistant. The decision support tool we investigate shall provide additional information during the final approach, hence for the part of the approach beyond the Final Approach Point (FAP) in case of a 3D approach or Final Approach Fix (FAF) in case of a 2D approach. Figure 25 and Figure 26 illustrate which tasks would be automated and performed by a functionality of the ML-constituent during the approach. Each functionality captured in the user stories is presented in a different color and described in more detail in Table 17.

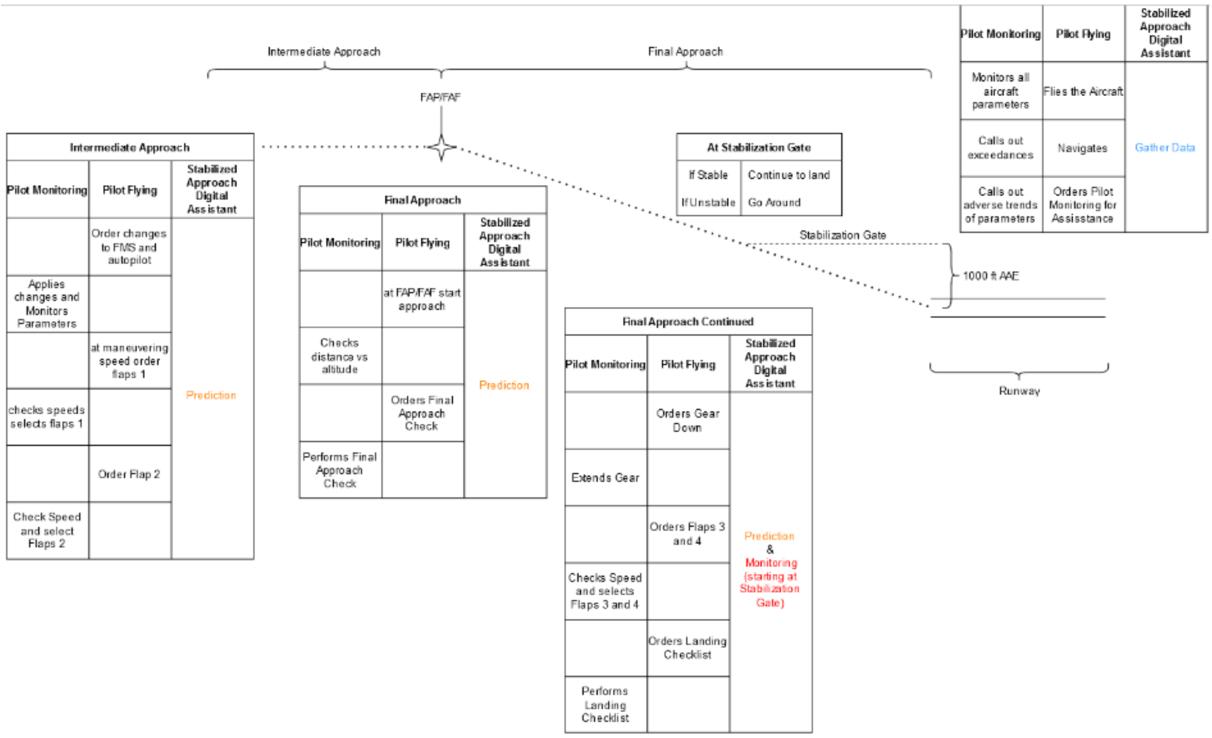


Figure 25 Sequential Task Analysis of the adopted system.

Additionally, the sequential task analysis adapts the HTA from 1. System Model, considering the additional tasks and new task distributions. The orange-colored boxes illustrate the Digital Assitant's tasks. The dark blue boxes illustrate tasks that both pilots perform, whereas the two colored blue boxes illustrate tasks performed by the Pilot Flying (PF) and the Auto Pilot (AP). The light blue boxes represent tasks performed by the Pilot Monitoring (PM), and Air Traffic Control-related tasks are illustrated in black. One can distinguish the two described modes, prediction and monitoring, as the Digital Assistant solely performs the prediction-related tasks. In contrast, the monitoring tasks are a functionality simultaneously performed by the PM and the Digital Assistant.

Funded by
the European Union

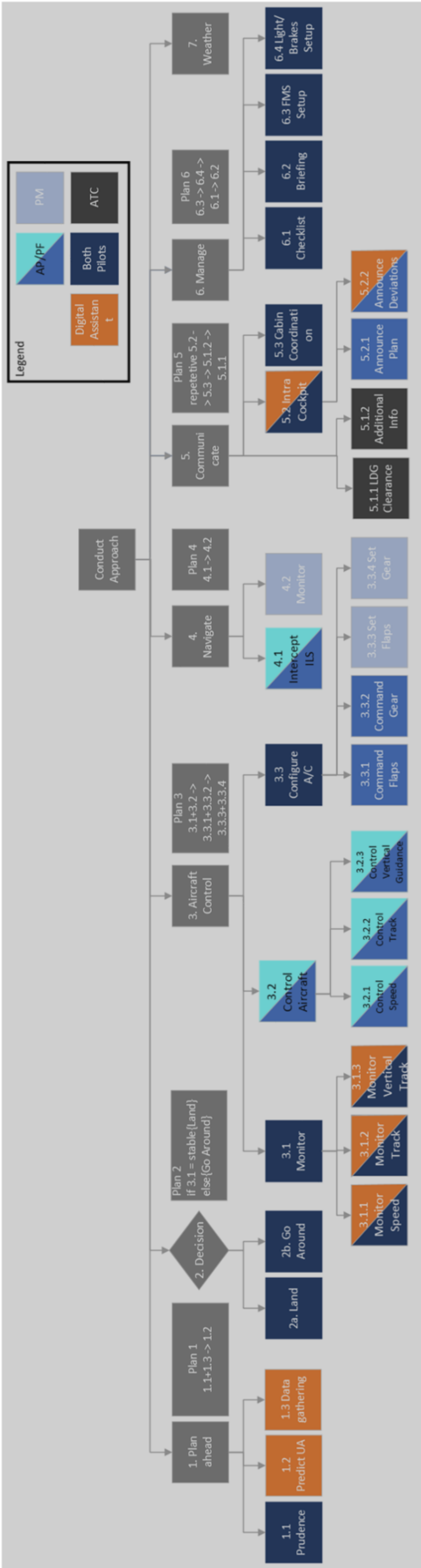


Figure 26 Hierarchical Task Analysis of the adopted system.

3.2.3 Human-Machine Interface Development

A well-designed Human-Machine Interface (HMI) is critical to ensuring a safe and effective digital assistant, especially in the high-workload environment of the approach phase of flight. Unclear, incomplete, or misleading information could lead to increased workload, high stress, and decreased situational awareness – potentially quickly making the in-flight situation more dangerous. Section 1.30 collects HMI Design Principles to guide the development of the stabilized approach digital assistant in terms of the HMI aspects to help ensure that the proposed design promotes increased safety in flight.

An initial diagram of the HMI-relevant elements commonly found on commercial aircraft is shown in Figure 15. This figure represents the design space for communicating output from the digital assistant to the pilot – via one of many combinations of visual, aural, and haptic cues.

Preferred HMI Elements

After discussions with various professional pilots, safety experts, and other interested parties from several airlines, the modified diagram shown in Figure 27 was generated. This second figure highlights the preferred elements of a well-designed stabilized approach digital assistant HMI. Items colored in green are very desirable elements of the digital assistant design; items in yellow may be helpful (but should first be evaluated for effectiveness during testing); grey items would likely not be beneficial for this digital assistant; and the one element in blue would be useful for system functionality information.

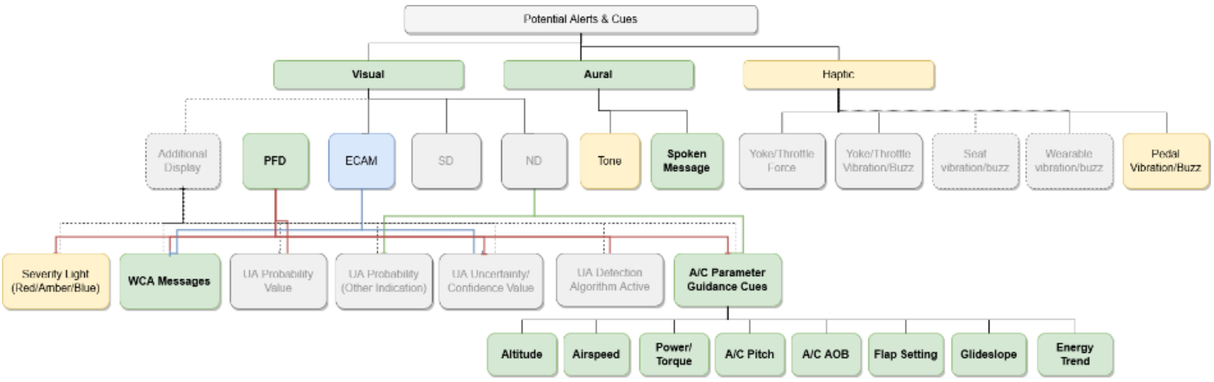


Figure 27 Preferred Stabilized Approach Digital Assistant HMI Design Elements.

Details about the research methodology, information requested, and preferences communicated are documented fully in Appendix B. Appendix B.8 HMI Questionnaire includes a particularly-detailed set of questions to enhance the research team's understanding of desired HMI design characteristics. The preferences expressed by pilots and airline representatives reflect a minimalist, high-TRL set of modifications and upgrades to existing cockpit systems. This is expected to help ensure the effectiveness and reliability of the system in a real airline operations context.

Recommended Stabilized Approach HMI Design

The PFD was identified as the primary display for digital assistant visual alerts and information, to ensure the pilot is not distracted away from the most critical parameters during an approach. On the PFD, Warning/Caution/Advisory (WCA) messages and visual modifications of relevant aircraft parameters (e.g. highlighting or reverse video) could efficiently and effectively display the primary findings and recommendations of the digital assistant to the crew. The ECAM could be used for system-related visual alerts (e.g. failure or override of the stabilized approach digital assistant system), as it already serves this function for other aircraft systems. No additional displays were recommended for displaying visual alerts related to this assistant, and a custom display was also considered undesirable.

Aural cues could separately alert the crew to predicted or current unstable approach conditions in the form of messages and tones. Some pilots suggested launching aural cues (e.g. short, spoken messages

Funded by the European Union

or a short tone) prior to visual alerts; this timing and mixture of potential message types should be evaluated in testing to determine the most effective approach. Otherwise, generally, aural cues would be launched simultaneously with appropriate visual cues on the PFD.

Haptic cues were unpopular overall within the group discussions for this digital assistant; however, stick forces and some vibration schemes may be included in testing to verify assumptions stated in the discussions.

It should be noted that HMI preferences were not uniform among all participants and research sources. Further information about the design preferences collected during discussions with potential users can be found in the User Stories, Section 3.3.1. While the diagram above and the "recommended" design description reflect the majority opinions, test cases should be designed to evaluate variations of interest around and outside the majority opinions.

3.3 Implementation

This section specifies what shall be implemented in the simulation environment within Unstable Approach Case Study Demonstration Task 4.2. Therefore, this section derives a set of high-level requirements from the user’s perspective as user stories, based on the high-level system description provided in sections System Model 3.1 and 3.2., as well as the findings from the user workshops of this task.

3.3.1 User Stories

In this section, we provide a selection of refined user stories extracted from the workshops. The original user stories, removed from each workshop, can be found in Appendix B - Workshops / Interviews, in each workshop section. Based on these refined user stories, we compare which functionalities and data sources are already accounted for in the discussed ML model and what has to be considered in the implementation phase in task 4.2 as a potential new feature. Additionally, some user stories contradict, which is natural when asking several people for opinions and feedback. In this case, we separate the opinions by employing User A / User B to capture the different opinions. Furthermore, we split the user stories into categories:

- objectives: capturing the envisioned impact from the users
- high-level functionality: capturing envisioned high-level functionalities from the users
- non-functional: capturing machine learning, data sources, and general non-functional related user input
- HMI: capturing the user's inputs regarding HMI

Objectives

Table 16 Objectives extracted from User/Stakeholder Workshops.

User Story ID	As a...	I want...	so that... / in order to...	Derived from...
O.1	Airline	that the Stabilized Approach Assistant reduces Unstable Approaches	increase safety in operation and avoid avoidable go-arounds.	WS1.O.1
O.2	Airline	that the Stabilized Approach Assistant reduces Unstable Approaches	provide a time and fuel benefit for the operation	WS3.O.2

O.3	Airline	that the Stabilized Approach Assistant increases situational awareness and the willingness to perform a go-around	increase safety in operation and the GA/UA ratio in case of an unstable approach.	WS1.O.2, WS3.O.1

High Level Functionality

Table 17 High Level Functionality User Stories extracted from Workshops.

User Story ID	As a...	I want...	so that... / in order to...	Derived from...
HL.P.1	Pilot	to be provided with information about the risk of an unstable approach and the associated contributing factors during the approach phase	have decision support on how to fly a stabilized approach.	WS1.HL.1, WS5.HL.1
HL.P.2	Pilot	to be provided with information about the associated contributing factors, in case of an unstable approach prediction, during the approach phase	have decision support on how to fly a stabilized approach.	WS1.HL.1, WS5.HL.1
HL.P.3	Pilot	the Stabilized Approach Assistant to indicate, if the approach situation can be stabilized at the stabilization gate	help me decide whether to continue the approach or not.	WS5.HL.2
HL.P.4	Pilot	*the Stabilized Approach Assistant to* provide optinal guidance on what task in the approach has to be performed until what point in the approach e.g. when to extract landing gear, speed brakes	I am stabilized at the stabilization gate.	WS3.HL.1, WS4.HL.1
Rational HL.P.4: In WS1.HMI.1, guidance was explicitly not wanted, in order to maintain manual flying skills				
HL.M.1	*Pilot*	*the Stabilized Approach Assistant* to monitor the approach after the stabilization gate and detect unstable approaches	there is an objective instance assisting the cockpit crew with go-around decision-making.	WS1.HL.2, WS2.HL.1
HL.P.5	Pilot	the possibility to turn off (standby) the Stabilized Approach Assistant complete or partially	it doesn't distract me in case of nuisance alerts.	WS5.HMI.2

Non-Functional

Table 18 Non-Functional User Stories extracted from Workshops.

User Story ID	As a...	I want...	so that... / in order to...	Derived from...
NF.1	Pilot	the Stabilized Approach Assistant to take into account: • time of day • experience of pilots (e.g. total flight hours and flight hours on aircraft type) • flight time	the ML model can take into account experience, fatigue, chaos level at arrival airport.	WS1.ML.1, WS4.ML.1

Funded by the European Union

		• cockpit composition • (long-time) stress of pilots • working atmosphere in the airline • the variation of aircraft tracks at a certain airport for predicting go-arounds		
NF.2	Pilot	the Stabilized Approach Assistant to be aware of airport-, airline- and aircraft-specific procedures (Operations Manual) e.g. (3D vs. 2D approaches)	the information computed by the assistant fits with the Standard Operating Procedures of my airline and aircraft.	WS1.ML.2
NF.3	Pilot	the Stabilized Approach Assistant to collect data for the predictions, starting with the intermediate approach phase	it takes into account more than the final approach phase, where the prediction is active.	
NF.4		the prediction threshold to be customizable	it can be configured according to the personal risk perception.	WS5.NF.1, WS1.HMI.2
Rational NF.4: According to WS5.NF.1 airlines want to minimize false negatives, whereas pilots want to minimize false positives. One way of resolving this contradiction would be to design the prediction threshold customizable, or optimize it according to the "area under curve" metric				
NF.5	Pilot	the Stabilized Approach Assistant to provide predictive information on the stability of the approach from the stabilization gate to landing, beginning at the final approach phase.	I can get a potentially unstable approach to be stable at and after the stabilization gate.	WS5.HL.2, WS5.NF.2 ,WS4.HL.2

HMI

Table 19 HMI User Stories extracted from Workshops and Questionnaire.

User Story ID	As a...	I want...	so that... / in order to...	Derived from...
HMI.1	Pilot (A,B,C)	the *Stabilized Approach Assistant* to provide easily understandable, intuitive information	prevent confusion and information overload	
HMI.2	Pilot (A)	the following information provided by the *Stabilized Approach Assistant* on the Primary Flight Display: indication of the stability trend of an approach primary factors contributing to UA prediction/actual situation unstable approach prediction probability	Maintain focus on the central PFD display Efficiently indicate problematic parameters	WS3.HMI.3, WS2.HMI.1
HMI.3	Pilot (B,C)	the following information provided by the *Stabilized Approach Assistant* on the Primary Flight Display:	Maintain focus on the central PFD display Efficiently indicate problematic parameters	

Funded by the European Union

		Warning/caution/advisory related to predicted or current unstable approach situation primary factors contributing to UA prediction/actual situation	Prevent information overload	
HMI.4	Pilot (A,B,C)	alerts provided by *the Stabilized Approach Assistant* to be presented in a more subtle way for UA predictions and a more prominent way for actual UA detection.	properly prioritize alerts based on the criticality of the information	WS3.HMI.4
HMI.5	Pilot (B,C)	alerts provided in a more prominent way (e.g. "Warning" level) for severe UA predictions	properly prioritize alerts based on the criticality of the situation	
HMI.6	Pilot (A,B,C)	the *Stabilized Approach Assistant* to integrate seamlessly into the existing cockpit systems Visual alerts only on the PFD One alert at a time for this system (most severe takes priority) Proper prioritization of visual & aural indications	minimize distraction and interference with other systems	WS5.HMI.1
HMI.7	Pilot (A)	the criticality level of the Stabilized Approach Assistant's indication to be dependent on prediction probability and the distance from the runway.	to indicate potential strategies for handling the situation.	WS4.HMI.1, WS4.HMI.2
HMI.8	Pilot (B,C)	the criticality level of indications to be dependent on the probability and severity of UA occuring	to enable earlier preparation for GA in severe cases	
HMI.9	Pilot (A,B,C)	the *Stabilized Approach Assistant* to not provide numbers or percentages for risk predictions	avoid complicated interpretations.	WS4.HMI.2
HMI.10	Pilot (A)	the *Stabilized Approach Assistant* to not provide aural information	avoid overlay with various existing aural warnings/cautions.	
HMI.11	Pilot (B,C)	the *Stabilized Approach Assistant* to provide aural cues to indicate an expected or current UA: Short, clear spoken messages (2-3 words) Messages include the offending parameter(s) & direction of issue (e.g. "Unstable - Airspeed Low") Minimal repetition Volume / style similar to TAWS Short tone before aural alerts Integrated into existing aural alert system, with proper prioritization	provide clear, concise, useful information & support regaining a stable approach profile	
HMI.12	Pilot (B)	the *Stabilized Approach Assistant* to indicate return to stability	clearly indicate improving condition	

Funded by the European Union

HMI.13	Pilot (C)	the *Stabilized Approach Assistant* to not indicate return to stability	prevent information overload	
HMI.14	Pilot (B,C)	the *Stabilized Approach Assistant* to provide the capability for its alerts to be muted	reduce distraction and masking of other alerts	
HMI.15	Pilot (B,C)	the *Stabilized Approach Assistant* to provide the capability for user override/disabling of the system	prevent false positives; properly handle system errors and other special situations	
HMI.16	Pilot (B,C)	the *Stabilized Approach Assistant* to announce its system health: Display digital assistant system errors, failures, and overrides on the ECAM Record digital assistant system errors, failures, and overrides in non-volatile memory for post-flight reporting	inform crew that the digital assistant is operable or operating in a degraded mode & track situations where this occured in flight.	
HMI.17	Pilot (B,C)	the *Stabilized Approach Assistant* to record the following in non-volatile memory for post-flight analysis: UA predictions (probabilities of UA & time horizons) Contributing factors (parameters causing UA) algorithms used to detect UA & corresponding uncertainties/confidence levels alerts shown to the flight crew alerts and other information not shown/masked for any reason pilot actions (override, muting, configuration changes, flight control inputs)	allow for post-flight analysis of stabilized approach digital assistant; increase clarity and trust	

3.3.2 Risk Assessment

Based on the Hierarchical Task Analysis performed in the previous sections, the tabular task analysis is introduced in this section, providing the engineers with a tool to identify dependencies among different tasks and risks associated with each task. This gives the designer not only additional insight to the system, but also helps in understanding how changes to the current system will impact existing work flows.

To avoid cluttering, Table 20 only shows the affected and new tasks. In Appendix C, the complete table is appended for the interested reader. The cells in column “subtasks” highlighted in red are there to emphasize those tasks that are affected by the decision support tool we aim to implement. Based on the envisioned changes, an initial discussion of potential risks is performed. For each risk, mitigation strategies are defined, which serve as starting point for the development of safety requirements during the development and implementation phase in task 4.2.

Table 20 Tabular Task Analysis of all affected tasks.

Task	Subtask	Task type	New task *or* affected by new task	Task affected by	Task affects	Risks	Mitigation Strategies
3.1 Monitoring aircraft states	3.1.1 Monitor speed	Acquisition, Analysis	Affected	1.2, 1.1	5.2.2, 2, 1, 5.2.1, 3.2	Complacency of Pilot Monitoring, Too strict UA limits will lead to nuisance alerts, Too strict UA limits will lead to unnecessary GA,	Recurring pilot training Appropriate setting of alert limits
	3.1.2 Monitor track	Acquisition, Analysis	Affected	1.2, 1.1	5.2.2, 2, 1, 5.2.1, 3.2	Complacency of Pilot Monitoring Too strict UA limits will lead to nuisance alerts, Too strict UA limits will lead to unnecessary GA,	Recurring pilot training Appropriate setting of alert limits
	3.1.3 Monitor vertical track	Acquisition, Analysis	Affected	1.2, 1.1	5.2.2, 2, 1, 5.2.1, 3.2	Complacency of Pilot Monitoring Too strict UA limits will lead to nuisance alerts, Too strict UA limits will lead to unnecessary GA,	Recurring pilot training Appropriate setting of alert limits
5.2 Intra-Cockpit	5.2.1 Announce plan to other pilot	Action	Affected	1, 4, 6		Complacency of Pilot Monitoring Misleading, unclear communication	Training Use of standard terminology
	5.2.2 Announce deviations	Action	Affected	3.1		Complacency of Pilot Monitoring Too strict UA limits will lead to nuisance alerts, Too strict UA limits will lead to unnecessary GA,	Training Use of standard terminology
1.1 Prudence		Analysis	Affected	1.2, 3		Complacency	Training
1.2 Predict UA		Analysis	New	3.2, 7, 6.3, 3.3, 3.2	3.2, 5.2, 3.1, 2, 3.3, 6.3	False predictions False positive - unnecessary GA → introduces another risk	

						False negative - suggests a stable approach even though it might be(come) unstable, situation perceived as safe even though there is an inherent risk	
1.3 Data Gathering		Acquisition	New				

Funded by the European Union

3.3.3 Simulation Environment

The Stabilized Approach Digital Assistant Case Study will be implemented in Task 4.2 in the Research Simulator at the Institute for Flight System Dynamics at Technical University of Munich. Since we have complete access to the flight dynamics model and the data streams within the simulator, we can integrate software products like digital assistants. The outputs of the digital assistant will be displayed to the pilots in the cockpit to provide prediction results. Several configurations of the display can be tested thanks to the flexibility in arranging items to be shown on the cockpit screens. In this section, we provide a brief overview of this simulation environment. The research simulator is a self-designed and self-build simulator, consisting of several subsystems. Figure 28 illustrates the DO-728, which is the simulated aircraft.

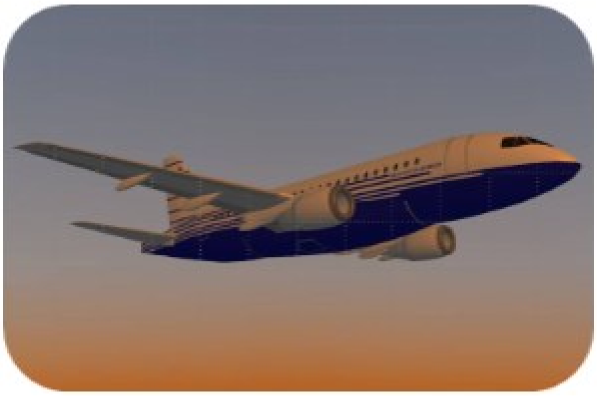


Figure 28: Do-728 Jet, Simulated in the Simulation Environment

Flight Dynamics Model

The flight dynamics model is the core model of the simulator. It is a high-fidelity simulation model, based on real-world tested aerodynamics and propulsion test data and implemented in Simulink by the Institute of Flight System Dynamics. Since the implementation is done in-house, the project has complete access and control over the simulation environment. Figure 29 exemplarily illustrates an simulation model in Simulink.

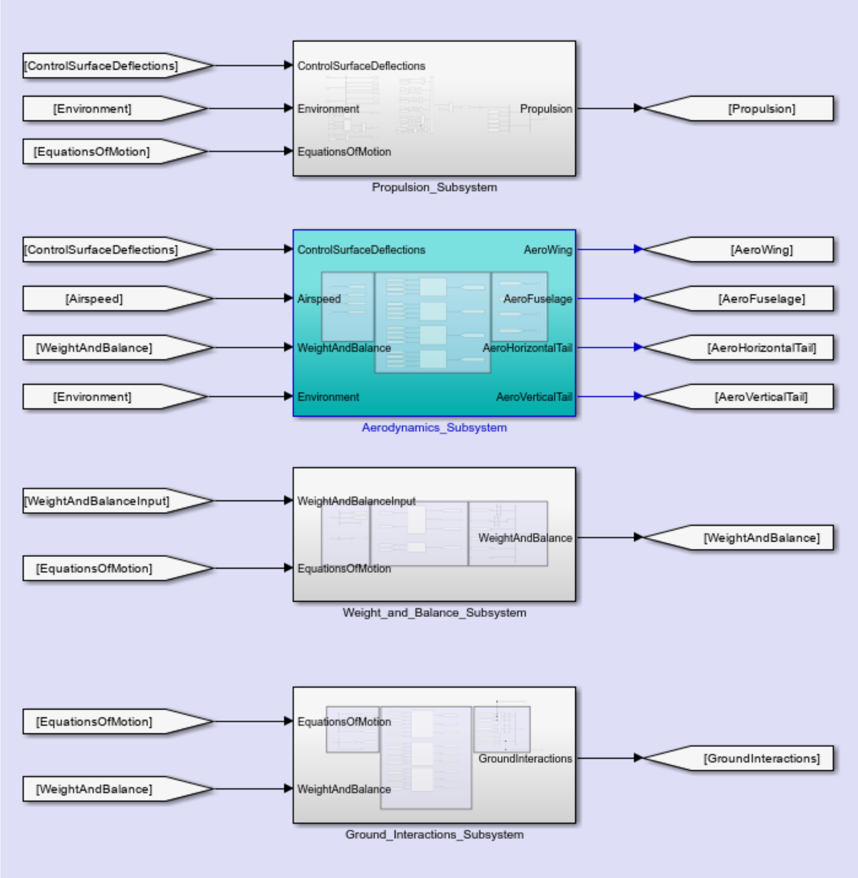


Figure 29: Exemplarily Simulator Model in Simulink

Flight Controls

The simulator provides common autopilot and auto-thrust modes and can be set up for different scenarios very quickly.

HMI

The simulator provides a generic cockpit, comparable to an A320 cockpit arrangement, with a fully customizable Primary Flight Display as well as a Navigation Display. Furthermore, aural information can be customized and played via speakers, similar to e.g. Terrain Awareness Warning System callouts. Figure 30 illustrates the HMI of the research simulator.

Figure 30: HMI of the Research Simulator

Visualization

For visualization, the Flight Gear software is used. The information of the Flight Dynamics Model are sent via UDP from the simulation PC to 3 visualization PC's providing a 180° vision from the cockpit.

3.3.4 Case Study Example Scenario

Table 21 provides an overview of Scenarios, planned to be used in task 4.2 for evaluation purposes.

Table 21: Case Study Scenarios

Scenario ID	Aircraft Type	Approach Type	Airport/Runway
Scen.1.3D	DO-728	3D - ILS	Antalya RWY 36R
Scen.1.2D	DO-728	2D - NDB	Antalya RWY 36R
Scen.2.3D	D728	3D - ILS	Sabia Gokcen
Scen.2.2D	D728	2D - NDB	Sabia Gokcen

Funded by
the European Union

References

[1] J. Scheiderer und H.-J. Ebermann, Human Factors im Cockpit, Berlin, Heidelberg: Springer Berlin Heidelberg, 2010.

[2] International Air Transport Association, „Safety Report," 2019. [Online]. Available: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwjM3Z2fkej_AhV_VfEDHYHEASMQFnoECBYQAQ&url=https%3A%2F%2Fwww.iata.org%2Fcontentassets%2F4d18cb077c5e419b8a888d387a50c638%2Fiata-safety-report-2019.pdf&usg=AOvVaw0pGtdru. [Zugriff am 29 06 2023].

[3] L. Li, S. Das, R. John Hansman, R. Palacios und A. N. Srivastava, „Analysis of Flight Data Using Clustering Techniques for Detecting Abnormal Operations," *Journal of Aerospace Information Systems,* Bd. 12, Nr. 9, p. 587–598, 2015.

[4] Z. Wang, L. Sherry und J. Shortle, „Feasibility of using historical flight track data to nowcast unstable approaches," IEEE, pp. 4C1-1-4C1-7.

[5] D. Walter, „Development of an Unstable Approach Detection and Prediction Model," Technical University of Munich, Munich, 2020.

[6] SafeClouds.eu, *Predictive analytics and identification of unknown hazards.*

[7] T. Blajev und W. Curtis, „Go-Around Decision-Making and Execution Project," 03 2017. [Online]. Available: https://flightsafety.org/wp-content/uploads/2017/03/Go-around-study_final.pdf. [Zugriff am 29 06 2023].

[8] Z. Wang, L. Sherry und J. Shortle, „Improving the Nowcast of Unstable Approaches," 2016.

[9] T. Jain, S. Misra und D. Truong, „Using Deep Learning to Predict Unstable Approaches for General Aviation Aircraft," *Journal of Aerospace Information Systems,* Bd. 19, Nr. 12, p. 811–817, 2022.

[10] SafeClouds.eu, *Historical Analysis through Decsriptive Analytics.*

[11] P. Hernandez, L. Beller, P. Koppitz und C. Argerich, „SafeOPS - Complete Data Pipeline Description and Machine Learning Solution," 18 05 2022. [Online]. Available: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5ec5dac75&appId=PPGMS. [Zugriff am 30 09 2022].

[12] P. Hernandez, C. Argerich und L. Beller, „SafeOPS - Human Interpretability Framework for the Selected User Stories," 10 08 2022. [Online]. Available: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5efe602ee&appId=PPGMS. [Zugriff am 30 09 2022].

[13] C. E. Billings und NASA, „Human-Centered Aviation Automation: Principles and Guidelines," 1996. [Online]. Available: https://ntrs.nasa.gov/api/citations/19960016374/downloads/19960016374.pdf.

[14] N. Minaskan, C. Alban-Dromoy, A. Pagani, J.-M. Andre und D. Stricker, „Human Intelligent Machine Teaming in Single Pilot Operation: A Case Study,“ in *Augmented Cognition*, Bd. 13310, Cham, Springer International Publishing, 2022, p. 348–360.

[15] F. Li und K. Chen, „High-Level Review Principles for Human-Machine Interface Design of Civil Aircraft Flight Deck,“ in *Advances in Human Aspects of Transportation*, Bd. 1212, Cham, Springer International Publishing, 2020, p. 624–631.

[16] G. Soudain, „EASA Concept Paper: First Usable Guidance for Level 1&2 Machine Learning Applications,“ 02 2023. [Online]. Available: https://www.easa.europa.eu/en/downloads/137631/en. [Zugriff am 29 06 2023].

[17] Boeing, [Online]. Available: https://www.boeing.com/commercial/.

[18] FlyByWire Simulations Documentation, „A320 Flight Deck Overview,“ [Online]. Available: https://docs.flybywiresim.com/pilots-corner/a32nx-briefing/flight-deck/. [Zugriff am 29 06 2023].

[19] W. Rouwhorst, R. Verhoeven, M. Suijkerbuijk, T. Bos, A. Maij, M. Vermaat und R. Arents, „Use of touch screen display applications for aircraft flight control,“ IEEE, p. 1–10.

[20] NAVBLUE S.A.S , „Runway Overrun Prevention System (ROPS+).,“ [Online]. Available: https://www.navblue.aero/product/rops-plus/#.

[21] Airbus S.A.S, „The Runway Overrun Prevention System,“ [Online]. Available: https://safetyfirst.airbus.com/app/themes/mh_newsdesk/documents/archives/the-runway-overrun-prevention-system.pdf.

[22] Skybrary, „Product Description - SmartRunwayTM /SmartLanding TM,“ [Online]. Available: https://www.skybrary.aero/sites/default/files/bookshelf/1974.pdf.

[23] C. E. Billings, Aviation automation, Mahwah, N.J: Lawrence Erlbaum Associates Publishers, 1997, p. 355.

[24] D. Harris, Engineering Psychology and Cognitive Ergonomics, Bd. 5639, Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.

[25] A. H. Roscoe, Hrsg., The practical assessment of pilot workload, Bd. 282, Neuilly-sur-Seine: AGARD, 1987, p. 134.

Funded by the European Union

Appendix A Concept of Operation – Self Assessment

This section shall serve as self-assessment regarding a concept of operation. In the following, we list the targeted content for a ConOPS from EASA and the US Department of Justice and compare our Use Case Definition against these criteria.

US Department of Justice – Requested Content

- Introduction
  - Project Description
  - Overview of Envisioned System
  - Document References
  - Glossary
- Goals, Objectives, and rationale for the new System
- Work Processes to be automated/supported
- High-Level functional requirements
  - High-level features
  - Additional features
  - Requirements Traceability
- High-Level operational Requirements
  - Non-functional requirements
  - Deployment and Support Requirements
  - Configuration and Implementation
  - System Environment
- User Classes and Modes of Operation
  - Classes/Categories of Users
  - User Classes Mapped to Functional Features
  - Sample Operational Scenarios
- Impact Considerations
  - Operational and Organizational Impacts
  - Potential Risks and Issues

EASA – Requested Content

EASA Objective ID	Objective
Objective CO-04	The applicant should define and document the ConOps for the AI-based system, including the task allocation pattern between the end user(s) and the AI-based system. A focus should be put on the definition of the operational design domain (ODD) and on the capture of specific operational limitations and assumptions.

The anticipated means of compliance are formulated with the following requests: The ConOps should be described at the level of the AI-based system, where the human is expected to achieve a set of high-level tasks. The ConOps should consider:

- an end-user-centric operational description of the AI-based system;
- the list of potential end users identified under Objective CO-01;

- how the end users will interact with the AI-based system: this description should be driven by the task allocation pattern between the end user(s) and the AI-based system, further dividing the high-level tasks identified under Objective CO-02 in as many sub-tasks as necessary;
- the definition of the operational design domain (ODD), including the specific operating limitations and conditions appropriate to the proposed operation(s);
- descriptions of the operational scenarios in their ODD; and some already identified risks, associated mitigations, limitations and conditions on the AI-based system.

EASA Objective ID	Objective
Objective CO-05	The applicant should perform a functional analysis of the system. The functional analysis consists in identifying, proposing a break-down of the high-level function(s) into sub-function(s), allocating the sub-function(s) to the subsystem(s), AI/ML constituents and items in line with the architecture choices. The delineation between AI/ML item and non-AI/ML item it performed at this stage: at least one item is allocated with AI functions and is thus considered an AI/ML item.

Funded by the European Union

Appendix B Stakeholder and User Workshops/Interviews

B.1 Workshop Preparation Template

General

Questions	Answers
1. What problems do you associate with unstable approaches?	
2. What are the reasons for unstable approaches?	
3. Can unstable approaches be reduced in your operation?	
4. How would you expect a prediction tool to benefit the operation?	
5. How would you expect a prediction tool to adversely affect the operation?	
6. What would the impact of the tool have to be, in order to be a reasonable investment?	

Machine Learning

Questions	Answers
1. Which solution would you preferer?	
2. Would you rather have it work differently?	
3. Are there data sources not mentioned here, you deem relevant?	
4. From which point in approach should data be used for prediction?	
5. What features do you think should be extracted from the data?	
6. Do you consider high Precision or Recall more important?	
7. How good would the prediction have to be, to be acceptable?	
8. Do you think 30 seconds is a large enough prediction time frame?	
9. Do you think 4NM is a valid prediction point?	

Operation

Questions	Answers
1. Which information/guidance would you expect from this machine-learning artifact? • Anomaly detection? • Guidance on how to stabilize a potential UA?	
2. How would you incorporate this assistant in your approach / How would you change the Approach Procedure, given this information? • Are you agreeing with the Hierarchical Task Analysis for the reference approach? • How would the Hierarchical Task Analysis change for the solution approach?	
3. What different approach types should the tool consider?	
4. How should the assistant behave, if an approach is outside of the training data domain?	

HMI

Questions	Answers
1. Which level of information should be presented a. Deterministic information b. Probabilistic information? c. Contributing factors for prediction?	
2.How should the information be provided? 1. a. Visual? b. Aural? c. Haptic?	
3. How Critical is the information compared to other systems? (Warning/caution/advisory)	

Funded by
the European Union

B.2 Workshop 2023/04/27

Nomenclature:

- PF - Pilot Flying
- UA - Unstable Approach
- GA - Go Around

B.2.1 Semi-Structured Interview:

General discussion at the start

What is your initial feeling regarding this idea?

- A stabilized approach shall prevent runway excursion
- If the tool helps to decide, whether an approach will be stabilized it might be useful. ← **potential user story**
- i.e. the tool tells the pilot x-wind limit will not be exceeded
- A detection tool (automated pilot monitoring) could be useful. ← **alternative/additional feature: SmartRunway and SmartLanding from Honeywell is such a system.**
- Useful for new/inexperienced and fatigued pilots (neglecting data protection issues for now) ← **Interesting idea to use data from crew management. This will be tricky/not possible due to data protection**
- Potential interesting data to train an ML model on:
    - Time of day
    - experience
- flight time (Every pilot is allowed to work a specified amount of hours a day.) ← **This could be taken from FDM data, possible (new) feature**
- Cockpit composition
- assistant can make predictions earlier or later
- Look specifically at stable approaches
- Take into account different aircraft and airline operating manuals. ← **This will part of defining the OD for the system /case study (we will only focus on a smaller OD but definitely, different A/Cs, different airlines (SOPs) will have to be treated**
- Over-reliance on such a system, not replacing the pilots only assisting ← **Potential user story: Concern especially when thinking about guidance on how to land stable vs. providing information on causes of potential unstable approach reasons.**
- Use AI to teach pilots to think more in terms of probabilities
- **Greatest benefit in preparing a pilot for the unlikely event ← potential target metric in simulator studies. change in situational awareness w and w/o predictive information**
- Flight Schools ← **potential Stakeholder, we haven't considered yet**
- Drone operations are fully operated, how are they handling potential instability? is it even considered?

When is an approach unstable, i.e. one data point outside the envelope is not an UA → So when is it?

- Compare to a traffic light: no black and white but yellow in between
- Look at the time that a deviation from the norm is not counteracted by the pilot flying (PF)

Questions **Answers**

Funded by the European Union

1. What problems do you associate with unstable approaches?	If there is information during flight preparation giving a risk of UA could help prepare a pilot's mindset for a GA. ← **Not directly the application we are looking into but worth considering**
2. What are the reasons for unstable approaches?	
3. Can unstable approaches be reduced in your operation?	
4. How would you expect a prediction tool to benefit the operation?	
5. How would you expect a prediction tool to adversely affect the operation?	Too many false positive indications may cause the crew the distrust the system → Pilots will abandon it ← **Potential technical requirement / non-functional requirement, regarding prediction quality → precision vs. recall**
6. What would the impact of the tool have to be, in order to be a reasonable investment?	Evidence based training Ga/UA ratio increase ← **Possible Performance Metric for Evaluation in Simulator** Reduce UAs ← **Goal and Possible Performance Metric for Evaluation in Simulator** Increase Situational Awareness ← **Goal and Possible Metric for Simulator Case Study** In risk assessment go to places we haven't gone

Operations

- Comment on Graphic → Turboprops do not have green dot speed
- Add stabilization criteria to the picture for completeness

Questions	Answers
1. Which information/guidance would you expect from a predictive tool / unstable approach prediction?	Differentiate functionality above and below stabilization gate ← **alternative/additional feature: SmartRunway and SmartLanding from Honeywell is such a system. This could be expressed in different features or modes** Pilot monitoring function after stabilization gate Predictive functionality before.
2. How would you incorporate this assistant in your approach / How would you change the Approach Procedure, given this information?	Information on potential reasons for UA is helpful but no indication on how to fly the A/C necessary → pilot should know how to stabilize the aircraft ← **Decision Support rather than an automation tool**
3. What different approach types should the tool consider?	Differentiate between precision and non-precision approach ← **Different Modes / Operational Domains**
4. How should the assistant behave, if an approach is outside of the training data domain?	An indication that tool is (in-)active ← **HMI User Story**

B.2.2 User Stories extracted after Interview

User Story ID	As a...	I want...	so that... / in order to...
WS1.HL.1	*Pilot*	*the Stabilized Approach Assistant to help me decide, whether an approach will be stabilized (or not)*	*I can get a potentially unstable approach to be stable at and after the stabilization gate.*
WS1.HL.2	*Pilot*	*the Stabilized Approach Assistant* to monitor the approach after the stabilization gate and detect unstable approaches	there is an objective instance assisting the cockpit crew with go-around decision-making.
WS1.ML.1	Pilot	*the Stabilized Approach Assistant* to take into account: time of day experience of pilots (e.g. total flight hours and flight hours on aircraft type) flight time cockpit composition for predicting go-arounds	the ML model can take into account experience and fatigue.
WS1.ML.2	Pilot	*the Stabilized Approach Assistant* to be aware of airport-, airline- and aircraft-specific procedures (Operations Manual) e.g. (3D vs. 2D approaches)	the information computed by the assistant fits with the Standard Operating Procedures of my airline and aircraft.
WS1.HF.1	Pilot	*that the Stabilized Approach Assistant* provides predictions on the stability of an approach and information about the causes for the prediction, but no guidance on how to stabilize the approach	over-reliance on an assistance system does not occur and the system does not replace a pilot (single pilot operation).
WS1.HF.2	Pilot	that *Stabilized Approach Assistant to minimize false positive predictions*	avoid distrust and disturbance by the cockpit crew.
WS1.O.1	Airline	that the *Stabilized Approach Assistant reduces Unstable Approaches*	increase safety in operation and avoid avoidable go-arounds
WS1.O.2	Airline	that the *Stabilized Approach Assistant* increases situational awareness and the willingness to perform a go-around	increase safety in operation and increase the GA/UA ratio in case of an actual unstable approach.

B.3 Workshop 2023/04/28

B.3.1 Semi Structured Interview

General

What is your initial feeling wrt this idea?

- Assist Callouts of PM, make callout/indications if a value is exceeded **← alternative/additional feature: SmartRunway and SmartLanding from Honeywell is such a system.**
- In Some situations pilot might not be aware of what will happen in a few seconds → here this would be useful
- Needs to look ahead, i.e. extrapolate pilot inputs, windshear, etc.
- Show the trend that an approach becomes unstable **← Potential User Story HMI**
- Suitable for fatigued pilots after long days **← Include Flight Time in the Feature Selection Process (Nr. of Legs might be difficult due to data protection)**
- No matter of Experience: pilot monitoring → PF has a tendency of having issues with one parameter in approach → callouts → PF might be offended ← automation could be an objective instance
- Modern aircraft digital → do not feel the environmental inputs as directly as with older aircraft

Funded by the European Union

Questions	Answers
1.What problems do you associate with unstable approaches?	
2. What are the reasons for unstable approaches?	ILS Loc / GS deviations Speed 10-20kts above target speed Speed loss 1000 ft gate: Landing checklist: alles erledigt trends negative and not counteracted SOPs adhered to
3. Can unstable approaches be reduced in your operation?	
4. How would you expect a prediction tool to benefit the operation?	Every 6 month simulator AI only talks to PM → problem with pilot incapacity Neutralize some of the problems associated with Human Factors, i.e. if a Co-Pilot does not speak up to the captain
5. How would you expect a prediction tool to adversely affect the operation?	
6. What would the impact of the tool have to be, in order to be a reasonable investment?	

Operations

Questions	Answers
Which level of information should be presented Deterministic information Probabilistic information? Contributing factors for prediction?	no % of UA tendency
2.How should the information be provided? Visual? Aural? Haptic?	if aural short and concise preferably visual indication PFD: color Aircraft symbol/Artificial Horizon from blue->yellow->red as simple as possible
3. How Critical is the information compared to other systems? (Warning/caution/advisory)	before 1000 ft gate advisory, later caution

B.3.2 User Stories Extracted from Interview

User Story ID	As a...	I want...	so that... / in order to...
WS2.HL.1	*Pilot*	*the Stabilized Approach Assistant to assist the pilot monitoring with callouts*	*there is an objective instance assisting the cockpit crew with go-around decision-making.*
WS2.HMI.1	Pilot	an indication of the stability trend of an approach	react to a changing/dynamic situation.

WS2.HL.2	Pilot	help from a prediction tool, especially in situations where fatigue is high / after long days	I can avoid mistakes due to fatigue.

B.4 Workshop 2023/05/12

B.4.1 Semi-Structured Interview

General

What is your initial feeling wrt this idea?

- Great potential when an approach is still relatively at the beginning, certain factors not in view e.g. wind near the airport
- UA criterion is not hard (e.g. speed downward limited by stall, upward no real criterion), in this situation predictions could help decide if the approach will be successful.
- Accidents resulting from UA often result from fixation (to land the aircraft despite not being stable). In these situations, such a tool would be useful **← User Story / Metric for Simulator Exercises (GA/UA rate), similar to WS1.O.1**
- If the approach is shorter than first thought, it helps to be aware of the situation.
- Reduce the hierarchy threshold from CoPi to Captain **← User Story (already similar one from first workshop)**
- Set a limit, from when what has to happen
- How to get to the 1000 ft gate that you are stable there
- Forecast to airport and weather what should be done differently to be stabilized
- The system provides additional experience, energy-state monitoring
- Comparison of target approach trajectories with actual trajectories

Questions	Answers
1. What problems do you associate with unstable approaches?	
2. What are the reasons for unstable approaches?	Missed Configuration The approach is unexpectedly longer Sporty flying style ATC Instructions
3. Can unstable approaches be reduced in your operation?	
4. How would you expect a prediction tool to benefit the operation?	The company has a high stable approach rate, however, a very low go-around/unstable approach rate ← Increasing the go-around willingness/preparedness of pilots **← Possible Objective/Goal of this tool, Similar to WS1.O.2**
5. How would you expect a prediction tool to adversely affect the operation?	Too many false alerts or indications for insignificant events → Pilots will ignore it or be distracted **← Similar to User story WS1.HF.2** If its too operating effort is too high, it can be distractive
6. What would the impact of the tool have to be, in order to be a reasonable investment?	Time, Fuel savings **← Secondary Objective for the tool** Safety Benefit**← Major Objective for the tool**

Operational

Funded by the European Union

Questions	Answers
1. Which information/guidance would you expect from this? Anomaly detection? Guidance on how to stabilize a potential UA?	fly fast + speed brakes → no use if ATC specifies speed fly slow + flaps → no use if ATC specifies speed gear out → conservative approach → useful hint for tool Off path descend feature → specifies how to descend with speed brakes and or gear out ← **Possible feature / incl. in user story** between maximum and minimum points between which certain measures must be taken to be able to stabilize at the stabilization gate
2. How would you incorporate this assistant in your approach / How would you change the Approach Procedure, given this information?	Based on the presented Task analysis, the Configuration and Monitoring Tasks change ← **Action item, adapt task analysis illustrations** example NARIDAS: Riskmonitor in shipping, traffic lights green, yellow red
3. What different approach types should the tool consider?	
4. How should the assistant behave, if an approach is outside of the training data domain?	

HMI

Questions	Answers
Which level of information should be presented Deterministic information Probabilistic information? Contributing factors for prediction?	no integers or percentages Contrib. Fac. relevant if it is not clear what contributes to instability
2.How should the information be provided? Visual? Aural? Haptic?	is displayed as a tape after a certain point. flashes as soon as something gets out of hand PFD Height deposit point always green → color-code to infer energy state. At Aural too many other systems already babbling/beeping there. **Potential User Story**
3. How Critical is the information compared to other systems? (Warning/caution/advisory)	subtle indication for prediction at first ← **Potential User Story** distinction can something still be done, must something be done, is it too late to do something
4. Do you know of any other assistant systems and do you use them?	

Notes for Task Analysis

- Green dot speed → maneuvering speed rename
- flaps 2 before FAP

- HTA: "Monitor/Control Vertical Guidance"
- Manage: break setup auto-brake an/aus, light setup, parameter anpassen
- Weather: separate task in HTA

B.4.2 User Stories Extracted from Interview

User Story ID	As a(n)...	I want...	so that... / in order to...
WS3.O.1	Airline	*the Stabilized Approach Assistant to* reduce the hierarchy gap in the cockpit,	Pilot Monitoring does not hesitate to call out an unstable approach.
	Pilot	*the Stabilized Approach Assistant to* provide guidance on what task in the approach has to be performed until what point in the approach e.g. when to extract landing gear, speed brakes	I am stabilized at the stabilization gate.
	Pilot	the *Stabilized Approach Assistant to compare the targeted approach trajectory with the actual trajectory*	provide information on deviations to the optimum.
WS3.O.2	Airline	the *Stabilized Approach Assistant* to provide a safety, time, and fuel benefit by minimizing unstable approaches	it is a reasonable investment.
WS3.HMI.1	Pilot	the *Stabilized Approach Assistant* to not provide numbers or percentages for risk predictions	avoid complicated interpretations.
WS3.HMI.2	Pilot	the *Stabilized Approach Assistant* to not provide aural information	avoid overlay with various existing aural warnings/cautions.
WS3.HMI.3	Pilot	the informa*tion provided by Stabilized Approach Assistant to be presented in the Primary Flight Display*	it is available on a central display
WS3.HMI.4	Pilot	the informa*tion provided by Stabilized Approach Assistant to be presented* in a subtle way, in case of predictions and a more prominent way in case of a detection	take into account the criticality of the information.

B.5 Workshop 2023/05/15

B.5.1 Semi Structured Interview

General

Questions	Answers
What problems do you associate with unstable approaches?	Runway overrun Hard landing More generally, landing-related incidents
2. What are the reasons for unstable approaches?	
3. Can unstable approaches be reduced in your operation?	
4. How would you expect a prediction tool to benefit the operation?	Prediction: less UA ← **Covered by Objective WS1.O.1**

Funded by the European Union

	UA indication: more G/A ← **LB: this and the point above imply two functionalities. One predictive and one reactive functionality. ← Covered by Objective WS1.O.2** Descend planning support increase the G/A willingness/readiness **← Covered by Objective WS1.O.2**
5. How would you expect a prediction tool to adversely affect the operation?	False negatives, nuisance alerts ← **WS1.HF.2** Overreliance on technique / Loss of manual flying skills **← WS1.HF.1 (overreliance) Add Loss of Manual Flying Skill as neg. Impact**
6. What would the impact of the tool have to be, in order to be a reasonable investment?	In non-normal operation: guidance tool which guides you on Runway **← Potential Feature** (XAVION) Benefit over existing Systemen (E-GPWS, ...) **← Impact User Story**

Machine Learning

Questions	**Answers**
1. Which solution would you preferer?	Both should be tested in the simulator
2. Would you rather have it work differently? A deterministic detection for example?	
3. Are there data sources not mentioned here, you deem relevant?	Soft factors: **← already covered by WS1.ML.1** Duration of Flight Shorthaul: Duty time Tracks of all aircraft at a certain airport **← potential feature** e.g.: Thunderstorm in Denver ⟶ ATCO performs Weather Avoidance e.g.: Thunderstorm in Frankfurt ⟶ little bit of chaos e.g.: Thunderstorm in Antalya ⟶ complete chaos Long-time stress **← possible addition to WS1.ML.1** Working atmosphere **← possible addition to WS1.ML.1**
4. From which point in approach should data be used for prediction?	From Top of Descent, data can be used for predictions start to compute prediction only make sense, when go-arounds make sense e.g. when the radio altimeter is active (2500 ft) Glide Intercept ⟶ maybe a few miles earlier
5. What features do you think should be extracted from the data?	
6. Do you consider a high Precision or Recall more important?	
7. How good would the prediction have to be, to be acceptable?	
8. Do you think 30 seconds is long enough prediction time frame?	7 sec for engine from IDLE to G/A Power → 30 sec long enough Airbus Trend Vector 10 sec Vorlauf um sich auf G/A vorbereiten → Ampel gelb Unterscheidung von "jetzt muss was passieren - rot" zu noch "kurz bedenken - gelb"

9. Do you think 4NM is a valid prediction point?

Operations

Questions	Answers
1. Which information/guidance would you expect from this?	No guidance Prediction and detection
2. How would you incorporate this assistant in your approach / How would you change the Approach Procedure, given this information?	Continuous risk assessment from begin of descend until leaving the runway The tool goes from info source to caution probabilistically based → later in the approach when accidents also become probable Key Card" - personal data with personal preference settings, risk-taking level ← **Potential ML/HMI User Story.** personalized risk estimation
3. What different approach types should the tool consider?	Precision- und Non-precision approaches (2- vs 3-D) und Manual Approach
4. How should the assistant behave, if an approach is outside of the training data domain?	

HMI

Questions	Answers
Which level of information should be presented Deterministic information Probabilistic information? Contributing factors for prediction?	
2.How should the information be provided? Visual? Aural? Haptic?	• traffic light system • green, below a certain threshold, • yellow, predicted above a certain threshold, • red, UA detected → no way of correction → go-around • PFD in the colored frame, compared to the Ambilight of TVs • Info too relevant for • ND • ECAM • blinking indication, when UA detected • In ECAM provide more details • Additional Callouts at the stabilization gate • band on the side of PFD representing 0-1 • Separating external and internal factors • external - cannot be influence by the pilot • No text, No number

Funded by the European Union

	• No additional display
3. How Critical is the information compared to other systems? (Warning/caution/advisory)	Advisory, Caution, Warning depending on criticality of prediction and distance to threshold ← **User Story (Implementation Idea, risk table with different thresholds and distances to runway defining color of indication)**
4. Do you know of any other assistant systems and do you use them?	

B.5.2 User Stories Extracted from Interview

User Story ID	As a(n)...	I want...	so that... / in order to...
WS4.Impact.1	Pilot	*the Stabilized Approach Assistant to* NOT cause loss of manual flying skills	pilots can still fly aircraft without the help of assisting systems.
WS4.Impact.2	Airline	*the Stabilized Approach Assistant to* provide a benefit compared to existing systems like TAWS	be a reasonable investment.
WS4.Feature.1	Airline	*the Stabilized Approach Assistant to* provide guidance towards the runway in non-normal operation (e.g. like Xavion)	it is a reasonable investment.
WS4.ML.1	Pilot	the Stabilized Approach Assistant to consider the variation of aircraft tracks at a certain airport	measure the level of chaos at the airport.
WS4.ML.2 (addition to WS1.ML.1)	Pilot	the Stabilized Approach Assistant to take into account the (long-time) stress and working atmosphere in the airline	
WS4.HL.1	Pilot	the Stabilized Approach Assistant to provide predictions on the stability of an approach from the GS Interception point on	I can correct potentially unstable approaches early on.
WS4.ML.3	Pilot	the Stabilized Approach Assistant to take into account flight performance information from the top of descent onwards	it can evaluate the complete descent phase when making predictions.
WS4.HMI.1	Pilot,	the criticality level of Stabilized Approach Assistant's indication to be dependent on prediction probability and the distance from the runway.	to indicate potential strategies for handling the situation.
WS4.HMI.2	Pilot	the indication of predictive information to be visual (preferably in the PFD), following a continuous color grading.	it is not too prominent, compared to other important information

B.6 Workshop 2023/05/23-24 Project Internal Workshop

Attendees:

- Pegasus
- TUM
- INX
- CAA

B.6.1 Semi Structured Interview

General

What is your initial feeling wrt this idea?

- Awareness tool ← **Decision Support Tool User Story**
- Include Jeppesen Reference pages
    - standard practices for each airport (i.e., 160 kt until 6 NM)
- Interesting airports ← **Case Study Definition / Scenario / OD**
    - Antalya RWY 36 R and RWY 36 C
    - Sabia Gokcen
- Regulatory framework
    - Certification - tool will probably have to be certified
    - Bow Tie - providing risks and steps such that a AI tool can be implemented
        - Crew and operator training
        - Change management
    - Updating in-air or only on ground will impact certification

Questions	Answers
1. What problems do you associate with unstable approaches?	
2. What are the reasons for unstable approaches?	Late configuration 1000ft AAL - configuration altitude High energy state Late ATC instruction Pilot error ⟵ **Possible high-level metric in the simulator case study would need to be refined to specify what errors can occur. Possibly also implicitly defined through the task analyses, as every task, if done wrong/forgotten is a possible error.** Glide slope
3. Can unstable approaches be reduced in your operation?	
4. How would you expect a prediction tool to benefit the operation?	The most benefit is reducing GA ← **Objective WS1.O.1 / WS3.O.2** Time Money
5. How would you expect a prediction tool to adversely affect the operation?	Adverse effects ← **Objective WS4.O.1 / WS3.O.2** Loss of training/skills Complacency - over-reliance Startle effect
6. What would the impact of the tool have to be in order to be a reasonable investment?	

Machine Learning

Funded by the European Union

Questions	Answers
1. Which solution would you preferer?	
2. Would you rather have it work differently?	
3. Are there data sources not mentioned here, you deem relevant?	
4. From which point in approach should data be used for prediction?	
5. What features do you think should be extracted from the data?	
6. Do you consider a high Precision or Recall more important?	
7. How good would the prediction have to be, to be acceptable?	
8. Do you think 30 seconds is large enough prediction time frame?	Prediction horizon: 30-60 sec before becoming unstable ⟵ **Operational User Story**
9. Do you think 4NM is a valid prediction point?	Starting at 8-10 NM out ⟵ **FAP/FAF as boundary Operational User Story**

Operations

- Predictions are mostly interesting for the final approach **← User Story, interesting regarding the system boundaries, especially for the prediction phase.**

HMI

Questions	Answers
Do you have experience with other digital assistant systems? If so, which ones? (e.g. runway overrun protection (ROW/ROPS), overspeed protection, etc)	Airbus' ROPS, two versions of the system Wet or Dry selector know -> calculates and warns according conditions Older version warns "if wet, runway too short" Honeywell Smart Runway Smart Landing
- If you have experience with other digital assistants, what positive HMI aspects would you like see implemented in future assistants?	People are familiar with SAM → it makes sense to integrate into such system **← HMI user story / Seamless integration**
- If you have experience with other digital assistants, what negative HMI aspects would you like a new system to avoid?	Minimize false negatives **← ML Objective/Non-functional User Story, in Precision vs. Recall tradeoff** / **this should be clarified again, seems counter-intuitive and contradicts the minimizing nuicance alerts WS1.HMI.2 → possibly max. AUC as a target metric? Clarify with Pablo Hernandez (INX)** Avoid an additional system **← Seamless Integration in existing tools** No Haptic feedback
2. Which other systems/messages would have a similar level of priority as this unstable approach assistant?	

(ie what level of priority would you assign to this digital assistant system?)	
	Green -> Go Blue -> Something is developing Blinking -> severe, click to stop blinking Override mode → turn system standby **← HMI feature** Similar to inhibit button in ROPS Differentiate scenarios for prediction Start algorithm/announcements at FAP Can approach still be stabilized by the 1000 ft? Is the approach passed the 1000 ft gate? → detection

B.6.2 User Stories Extracted from Interview

User Story ID	As a(n)...	I want...	so that... / in order to...
WS5.HL.1	Pilot	to be provided with information about the risk of an unstable approach and the associated contributing factors during the approach phase	have decision support on how to fly a stabilized approach.
WS5.HL.2	Pilot	the stabilized approach digital assistant to indicate, if the approach situation	
WS5.HL.2	Pilot	the stabilized approach digital assistant to provide predictive information on the stability of the approach from the 1000ft gate to landing, beginning at the final approach phase.	*I can get a potentially unstable approach to be stable at and after the stabilization gate.*
WS5.HMI.1	Pilot	the stabilized approach digital assistant to integrate seamlessly into the existing HMI concept and similar warning systems	it minimizes distraction and provides intuitive handling.
WS5.HMI.2	Pilot	the possibility to turn off (standby) the stabilized approach digital assistant complete or partially	it doesn't distract me in case of nuisance alerts.
WS5.NF.1	Airline	the stabilized approach digital assistant to be configured towards minimizing false negatives	it covers most unstable approaches.
WS5.NF.2	Pilot	the stabilized approach digital assistant to provide predictions from the final approach point / fix on	*I can get a potentially unstable approach to be stable at and after the stabilization gate.*
WS5.Scenario.1	Airline	the case study to include: Antalya RWY 36 R and RWY 36 C Sabia Gokcen	test the idea in a relevant environment.

B.7 Workshop 2023/05/26

B.7.1 Semi-Structured Interview

General

What is your initial feeling wrt this idea?

- Added value and safety gains **← Impact**

Funded by the European Union

- Redundancy is further increased
- 1000 ft stabilized call → triggered by system (visual) ← **WS1.HL.2**
- Probabilistic display, color coded.
- Says at gate "continue" or "go around" ← **WS1.HL.2**
- AI → recommendation character ← **WS5.HL.1**
- Ground speed mini function
  - ATIS moderate turbulence to Malaga
  - speed self selected
  - hard to be stabilized at 1000 ft
  - at special airports with special winds (Funchal) after 1000 ft gate over/under speed will be neglected ← **WS1.ML.2**

Questions	Answers
1.What problems do you associate with unstable approaches?	UA starts already in cruise or early approach Stabilized approach = stabilized at the gate
2. What are the reasons for unstable approaches?	Descent planning Environmental factors ATC Work in Cockpit, skills
3. Can unstable approaches be reduced in your operation?	
4. How would you expect a prediction tool to benefit the operation?	
5. How would you expect a prediction tool to adversely affect the operation?	
6. What would the impact of the tool have to be, in order to be a reasonable investment?	Information of probability of go around at a given airport at a given time is valuable This awareness is a benefit to safety

Machine Learning

Questions	Answers
1. Which solution would you preferer?	continuously solution preferred
2. Would you rather have it work differently? A deterministic detection for example?	
3. Do you know if there are data sources not mentioned here, you deem relevant?	
4. From which point in approach should data be used for prediction?	
5. What features do you think should be extracted from the data?	
6. Do you consider a high Precision or Recall more important?	
7. How good would the prediction have to be, to be acceptable?	
8. Do you think 30 seconds is large enough prediction time frame?	

9. Do you think 4NM is a valid prediction point?	

Operational

Questions	Answers
1. Which information / guidance would you expect from this?	
2. How would you incorporate this assistant in your approach / How would you change the Approach Procedure, given this information?	
3. What different approach types should the tool consider?	Matrix for approach types Which approach has the most go-arounds?
4. How should the assistant behave, if an approach is outside of the training data domain?	

- Special airports are particularly important, e.g. **← Interesting Case Studies but not sure if possible within this project**
  - Funchal
  - Heraklo

Modify task analysis

- Setting Go Around altitude → check distance vs. altitude
- Intermidiate approach column→ start monitor
- FAP → start giving information
-
  - Follow ILS → pilot flying
  - Monitor speed pilot monitoring → both Piloten
  - AI shoudl do:
    - monitor
    - Communicate → announce deviations
    - Manage → FMS setup
    - Weather → monitoring wind speeds

HMI

Questions	Answers
1. Which level of information should be presented a. Deterministic information b. Probabilistic information? c. Contributing factors for prediction?	• What is default status if everything is ok ○ silence or output • Contributing factors • How to announce system is not working correctly ○ FMA - AI Anzeige
2.How should the information be provided? 1.	• In field of vision ○ PFD

Funded by the European Union

a. Visual? b. Aural? c. Haptic?	○ ND • ECAM • Only indication for PM?
3. How Critical is the information compared to other systems? (Warning/caution/advisory)	

B.7.2 User Stories Extracted from Interview

User Story ID	As a...	I want...	so that... / in order to...
WS6.HIMI.1	Pilot	the PFD, ND to provide the information of the stabilized approach digital assistant and the ECAM to provide error messages regarding the stabilized approach digital assistant	the important information are within the file of vision

B.8 HMI Questionnaire

1. Information Requested

The full list of information requested in the HMI questionnaire is included below.

Question/Information Requested
Email Address
Your Name (LAST NAME, first name; or group name)
Current Job Title(s) / Role(s)
Current Employer(s) & Affiliation(s)
Please select all that apply for you: __ Licensed fixed-wing pilot (any rating) __ Licensed rotary-wing pilot (any rating) __ Commercial pilot, currently flying for an airline. __ Commercial pilot, flying for a non-airline entity. __ Employee working in a safety-focused group __ Employee in a design/engineering-focused group. __ Other...
How many hours do you have in the following aircraft? [Airbus A320 Series]
How many hours do you have in the following aircraft? [Airbus A321 Series]
How many hours do you have in the following aircraft? [Boeing 737 Series (not MAX)]
How many hours do you have in the following aircraft? [Boeing 737 MAX Series]
How many hours do you have in the following aircraft? [Other (please write in the following box)]
"Other" aircraft - Please describe here
Do you have any prior experience with digital assistants in aviation?
If Yes, which digital assistants do you have experience with (e.g. Runway overrun protection)?
(If yes) Please describe positive HMI aspects you would like see implemented in future assistants.

Question/Information Requested
(If yes) Please describe negative HMI aspects you would like avoided in future assistants.
The following shows potential options for visual, aural, and haptic cues. Please comment if you see anything missing / incorrect / inadvisable.
Are there any new HMI technologies/capabilities that you would like to see incorporated into this digital assistant?
Is there anything that can be added to the HMI that would help you to better "trust" the information provided by the digital assistant? (i.e. what additional in-flight or post-flight information would help provide transparency & contribute to pilot trust of the digital assistant?)
Is there any information about the digital assistant algorithms should be shown alongside WCAs/guidance markers (in-flight)? If so, how/where? (e.g. uncertainty/confidence, algorithms used, number of occurences, time until unstable/stable, likelihood/degree of instability, etc)
Do you feel that an overall "override" feature should be available for the digital assistant system?
(If yes) Do you have any specific suggestions regarding the design/inclusion of an override feature? (e.g. which existing button could be used, how the override could work, etc)
Please choose the level of visual WCA/alert severity that you think is most appropriate for the following situations: [UA likely later (u_long > 90%)]
Please choose the level of visual WCA/alert severity that you think is most appropriate for the following situations: [UA expected soon (u_short > 90%)]
Please choose the level of visual WCA/alert severity that you think is most appropriate for the following situations: [UA detected (u_now > 95%)]
Please choose the level of visual WCA/alert severity that you think is most appropriate for the following situations: [Stability Re-established]
Please choose the level of visual WCA/alert severity that you think is most appropriate for the following situations: [High severity UA Very likely / Detected]
Please choose the level of visual WCA/alert severity that you think is most appropriate for the following situations: ["Other" (Please provide below)]
"Other" situation/condition - please describe here.
How much would you expect the following visual cues to help the pilot regain/ensure stability in an unstable approach (1 = "not at all / unlikely"): [WCA "UNSTABLE APP" (or similar)]
How much would you expect the following visual cues to help the pilot regain/ensure stability in an unstable approach (1 = "not at all / unlikely"): [Static markers on A/C parms]
How much would you expect the following visual cues to help the pilot regain/ensure stability in an unstable approach (1 = "not at all / unlikely"): [Static markers - runway/nav display]

Funded by the European Union

Question/Information Requested
How much would you expect the following visual cues to help the pilot regain/ensure stability in an unstable approach (1 = "not at all / unlikely"): [Blinking/Flashing markers - A/C parms]
How much would you expect the following visual cues to help the pilot regain/ensure stability in an unstable approach (1 = "not at all / unlikely"): [Blinking/Flashing symbols - runway/nav display]
How much would you expect the following visual cues to help the pilot regain/ensure stability in an unstable approach (1 = "not at all / unlikely"): [Highlighted Parameter(s)]
How much would you expect the following visual cues to help the pilot regain/ensure stability in an unstable approach (1 = "not at all / unlikely"): ["Other" - Please describe below]
"Other" style of visual cues - Please describe here.
How much would you expect the following visual cues to increase pilot workload or cause confusion/distraction in an unstable approach (1 = "not at all / unlikely"): [WCA "UNSTABLE APP" (or similar)]
How much would you expect the following visual cues to increase pilot workload or cause confusion/distraction in an unstable approach (1 = "not at all / unlikely"): [Static markers on A/C parms]
How much would you expect the following visual cues to increase pilot workload or cause confusion/distraction in an unstable approach (1 = "not at all / unlikely"): [Static markers - runway/nav display]
How much would you expect the following visual cues to increase pilot workload or cause confusion/distraction in an unstable approach (1 = "not at all / unlikely"): [Blinking/Flashing markers - A/C parms]
How much would you expect the following visual cues to increase pilot workload or cause confusion/distraction in an unstable approach (1 = "not at all / unlikely"): [Blinking/Flashing symbols - runway/nav display]
How much would you expect the following visual cues to increase pilot workload or cause confusion/distraction in an unstable approach (1 = "not at all / unlikely"): [Highlighted Parameter(s)]
How much would you expect the following visual cues to increase pilot workload or cause confusion/distraction in an unstable approach (1 = "not at all / unlikely"): ["Other" - Please describe below]
"Other" style of visual cues - Please describe here.
Regarding the persistance of WCAs/symbols displayed for an unstable approach, do you have any specific preferences/guidance? i.e. Should WCA messages disappear automatically if the condition improves? Or should they change in some specific manner? Should a Warning overwrite a caution, or should both remain "latched"/visible?
Regarding the display style of WCAs/symbols shown for an unstable approach, do you have any specific preferences/guidance for the conditions listed above (e.g. UA likely later, UA likely soon, etc)?

Question/Information Requested
Are there particular screens or parameters that should not have any visual cues added to it?
Do you have any other ideas or suggestions regarding visual cues for this digital assistant?
How much would you expect the following aural cues to help the pilot regain/ensure stability in an approach (1 = "not at all / unlikely"): ["UNSTABLE" (repeated, when detected)]
How much would you expect the following aural cues to help the pilot regain/ensure stability in an approach (1 = "not at all / unlikely"): ["UNSTABLE in 30 seconds" (predicted)]
How much would you expect the following aural cues to help the pilot regain/ensure stability in an approach (1 = "not at all / unlikely"): ["STABLE in 10 SECONDS"]
How much would you expect the following aural cues to help the pilot regain/ensure stability in an approach (1 = "not at all / unlikely"): ["UNSTABLE - Airspeed"]
How much would you expect the following aural cues to help the pilot regain/ensure stability in an approach (1 = "not at all / unlikely"): ["Unstable - Airspeed low"]
How much would you expect the following aural cues to help the pilot regain/ensure stability in an approach (1 = "not at all / unlikely"): [STABLE (after UA not detected for ~10 sec)]
How much would you expect the following aural cues to help the pilot regain/ensure stability in an approach (1 = "not at all / unlikely"): [Long tone]
How much would you expect the following aural cues to help the pilot regain/ensure stability in an approach (1 = "not at all / unlikely"): [Short tone]
How much would you expect the following aural cues to help the pilot regain/ensure stability in an approach (1 = "not at all / unlikely"): ["Other" - Please describe below.]
"Other" - Please describe here
How much would you expect the following aural cues to increase pilot workload or cause confusion/distraction in an approach (1 = "not at all / unlikely"): ["UNSTABLE" (repeated, when detected)]
How much would you expect the following aural cues to increase pilot workload or cause confusion/distraction in an approach (1 = "not at all / unlikely"): ["UNSTABLE in 30 seconds" (predicted)]
How much would you expect the following aural cues to increase pilot workload or cause confusion/distraction in an approach (1 = "not at all / unlikely"): ["STABLE in 10 SECONDS"]
How much would you expect the following aural cues to increase pilot workload or cause confusion/distraction in an approach (1 = "not at all / unlikely"): ["UNSTABLE - Airspeed"]
How much would you expect the following aural cues to increase pilot workload or cause confusion/distraction in an approach (1 = "not at all / unlikely"): ["Unstable - Airspeed low"]
How much would you expect the following aural cues to increase pilot workload or cause confusion/distraction in an approach (1 = "not at all / unlikely"): [STABLE (after UA not detected for ~10 sec)]

Funded by the European Union

Question/Information Requested
How much would you expect the following aural cues to increase pilot workload or cause confusion/distraction in an approach (1 = "not at all / unlikely"): [Long tone]
How much would you expect the following aural cues to increase pilot workload or cause confusion/distraction in an approach (1 = "not at all / unlikely"): [Short tone]
How much would you expect the following aural cues to increase pilot workload or cause confusion/distraction in an approach (1 = "not at all / unlikely"): ["Other" (Please describe below)]
"Other" aural cues - Please describe here
Regarding the relative loudness (volume) of aural cues, do you have any specific preferences/guidance? (e.g. quieter than "engine out"/ louder than radar altitude call-outs)
Regarding the repetition of aural cues, do you have any specific preferences/guidance?
Do you have any other ideas or suggestions regarding aural cues for this digital assistant?
How much would you expect the following haptic cues to help the pilot regain/ensure stability in an approach (1 = "not at all / unlikely"): [Yoke/throttle force]
How much would you expect the following haptic cues to help the pilot regain/ensure stability in an approach (1 = "not at all / unlikely"): [Yoke/throttle vibe/buzz]
How much would you expect the following haptic cues to help the pilot regain/ensure stability in an approach (1 = "not at all / unlikely"): [Seat vibe/buzz]
How much would you expect the following haptic cues to help the pilot regain/ensure stability in an approach (1 = "not at all / unlikely"): [Wearable vibe/buzz]
How much would you expect the following haptic cues to increase pilot workload or cause confusion/distraction in an approach (1 = "not at all / unlikely"): [Yoke/throttle force]
How much would you expect the following haptic cues to increase pilot workload or cause confusion/distraction in an approach (1 = "not at all / unlikely"): [Yoke/throttle vibe/buzz]
How much would you expect the following haptic cues to increase pilot workload or cause confusion/distraction in an approach (1 = "not at all / unlikely"): [Seat vibe/buzz]
How much would you expect the following haptic cues to increase pilot workload or cause confusion/distraction in an approach (1 = "not at all / unlikely"): [Wearable vibe/buzz]
How much would you expect the following haptic cues to increase pilot workload or cause confusion/distraction in an approach (1 = "not at all / unlikely"): ["Other" (Please Describe below)]
"Other" haptic/tactile cues- Please describe here
Regarding the relative strength of haptic cues, do you have any specific preferences/guidance?
Regarding the repetition of haptic cues, do you have any specific preferences/guidance?
Do you have any other ideas or suggestions regarding haptic/tactile cues for this digital assistant?

Funded by
the European Union

2. Summary/Key Findings of Results

General Info:

5 total responses: 3 from Pegasus; 2 from CAA/CAAi; all have some DA experience

- 3 professional airline pilots (captains, 5000+ hrs)
- 1 commercial transport helicopter pilot (1000-2000 hrs)
- 1 PPL(A) / Flight test engineer (closer to 250 hrs)
- Previous DA experience:
    - nearly all know/use SR/SL (Honeywell), EGPWS, RAAS, EFB
    - one respondent mentioned synthetic vision trajectory predictors & CLP correction cues (for maintaining Nr)

UA DA HMI Should Include/Be:

- uncluttered, intuitive, clear, accurate, brief, harmonious with other cockpit announcements
- inhibit/override capability (e.g. using the EGPWS inhibit button(s))
- aural + visual cues (maybe not haptic; mixed results regarding haptic)
- record of UA DA alerts/cautions/decisions in aircraft "quick access" (or similar post-flight) recorders; ability to replay the situation later, including UA DA aspects
- easily-explainable to pilots in ground school

Visual Alert Preferences:

- Avoid visual clutter & distractions:
    - No additional info (e.g. algorithm used, uncertainty/confidence, etc)
    - Limit UA DA info to the PFD (except for UA DA system failure/inhibit information)
    - Brief text in WCA message (including text related to parm(s) triggering UA alert may be ok - should be tested)
    - UA DA system failure or inhibit would be best shown on the ECAM
- Avoid unclear coloring (i.e. avoid anything requiring "interpretation"/non-standard)
    - Colors/alerts should follow Honeywell HMI & related norms (e.g. Warnings: Red; Cautions: Amber; Advisories: Green; Information: Cyan/Blue)
- WCAs start with "expected" UA condition:
    - UA likely later (with >90% certainty): mixed results: 3 say "Caution"; 2 say "Advisory"
    - UA expected soon (with >90% certainty): "Caution" (all)
    - High severity UA very likely (or detected) → Warning (all)
    - UA detected (> 95% certainty UA condition=true) → Warning (all)
    - Stability regained (UA = false, with >95% certainty) → Advisory (all)
        - One option for showing this is to re-color the "UNSTABLE" WCA as cyan / blue → indicating transition out/regaining stability
- Indicating triggering parameter(s) directly would be useful (should be tested): Highlighting / drawing a box / changing color (need to test to see which is most effective)
- Persistence: some disagreement here; may depend on the user/pilot preferences

Funded by the European Union

    - Option A: Display 2 consecutive WCAs on PFD for transition to UA; 1 for regaining stability:
        - UA Expected
        - UA detected
        - Stable APP (flash twice, in cyan/blue)
    - Option B: WCA persists until Unstable expected/detected = false, then briefly re-color or replace with Information message indicating stability achieved
    - Option C: WCA persists until Unstable expected/detected = false, simply no message when stability regained
- Avoid unnecessary repetition:
    - Only show one UA DA message at a time
    - Warnings replace cautions; cautions replace advisories, etc (of same type)

<u>Aural Alert Preferences</u>

- Aural cues are expected to be very useful if correctly implemented:
    - Minimize length of message
    - Integrate properly with other messages (from other systems)
    - Muting feature available
- Aural Message Wording:
    - Favorite overall (for each parameter): "Unstable - Airspeed Low" or "Unstable - Airspeed" → should test
    - Mixed-results, but worth testing:
        - Announcing UA timing (e.g. "Unstable in 30 sec" / "Unstable in 10 sec")
            - Pegasus → high approval
            - others → very low approval
        - Announcing return to stability (e.g. "Stable")
        - Short tones
        - Giving "Go Around" direction
    - Things to avoid:
- Loudness/volume levels: Follow Windshear / TAWS styles
    - Warning → louder, like "Terrain - Pull Up"
    - Caution → slightly quieter
- Repetition: several ideas here
    - a) Warning → 2x in 5 sec intervals, repeating until untrue; Caution → 2x in 10 sec intervals, repeating until untrue
    - b) same as "a" except only 2 repetitions; then stop aural messages
    - c) only repeat Warning

<u>Haptic Cue Preferences</u>

- Haptic cues deemed undesirable for this digital assistant (by survey respondents)

Funded by
the European Union

- One survey participant suggested trying pedal kick/buzzes (like ABS brakes / anti-shimmying systems) → should test
- Boeing HF expert suggests stick forces → should test

Funded by
the European Union

Appendix C Tabular Task Analysis

Table 22: Complete Tabular Task Analysis

Task	Subtask	Task type	New task *or* affected by new task	Task affected by	Task affects	Risks	Design requirements	Comment
3.1 Monitoring aircraft states	3.1.1 Monitor speed	Acquisition, Analysis	Affected	1.2, 1.1	5.2.2, 2, 1, 5.2.1, 3.2	Complacency of Pilot Monitoring, Too strict UA limits will lead to nuisance alerts, Too strict UA limits will lead to unnecessary GA,	Recurring pilot training Appropriate setting of alert limits	No new or affected tasks
	3.1.2 Monitor track	Acquisition, Analysis	Affected	1.2, 1.1	5.2.2, 2, 1, 5.2.1, 3.2	Complacency of Pilot Monitoring Too strict UA limits will lead to nuisance alerts, Too strict UA limits will lead to unnecessary GA,	Recurring pilot training Appropriate setting of alert limits	
	3.1.3 Monitor vertical track	Acquisition, Analysis	Affected	1.2, 1.1	5.2.2, 2, 1, 5.2.1, 3.2	Complacency of Pilot Monitoring Too strict UA limits will lead to nuisance alerts, Too strict UA limits will lead to unnecessary GA,	Recurring pilot training Appropriate setting of alert limits	
3.2 Controlling the aircraft	3.2.1 Control speed	Action		3.1				
	3.2.2 Control track	Action		3.1				

	3.2.3 Control vertical track	Action		3.1				
3.3 Configuring aircraft	3.3.1 Command Flaps	Action		3.1				
	3.3.2 Command Gear	Action		3.1				
	3.3.3 Set Flaps	Action		3.1				
	3.3.4 Set Gear	Action		3.1				
4.1 Intercepting and follow ILS		Action						
4.2 Monitor		Analysis, Acquisition						
5.1 ATC	5.1.1 Issue landing clearance	Analysis, Action						
	5.1.2 Provide additional information (e.g. weather)	Analysis, Action						
5.2 Intra-Cockpit	5.2.1 Announce plan to other pilot	Action	Affected	1, 4, 6		Complacency of Pilot Monitoring Misleading, unclear communication	Training Use of standard terminology	

Funded by the European Union

	5.2.2 Announce deviations	Action	Affected	3.1		Complacency of Pilot Monitoring Too strict UA limits will lead to nuisance alerts, Too strict UA limits will lead to unnecessary GA,	Training Use of standard terminology	
5.3 Cabin coordination		Action		1.1				
6.1 Checklist		Action		1.1				
6.2 Briefing		Action		1.1				
6.3 FMS		Action		1.1				
6.4 Light/Brakes		Action		1.1				
1.1 Prudence		Analysis	Affected	1.2, 3		Complacency	Training	
1.2 Predict UA		Analysis	New	3.2, 7, 6.3, 3.3, 3.2	3.2, 5.2, 3.1, 2, 3.3, 6.3	False predictions False positive - unnecessary GA → introduces another risk False negative - suggests a stable approach even though it might be(come) unstable, situation perceived as safe even though there is an inherent risk		
1.3 Data Gathering		Acquisition	New					

Funded by the European Union