



Virtueller Crewassistent - Kontextspezifische Unterstützung für „Reduced Crew Operations“ mittels Künstlicher Intelligenz

Abschlussbericht 2024

Jonas Füllgraf, Christian Roxlau, Sven Bollmann
[j.fuellgraf] / [c.roxlau] / [s.bollmann] @tu-braunschweig.de

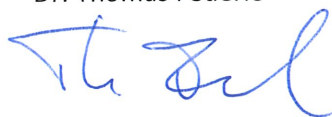
Zuwendungsempfänger: Technischen Universität Braunschweig,
Institut für Flugführung

Förderkennzeichen: 20D1910D

Laufzeit des Vorhabens: 01.07.2020 - 30.06.2024

Erstellungsdatum: 23.09.2024

Projektleitung: Dr. Thomas Feuerle



Inhaltsverzeichnis

Abkürzungsverzeichnis	3
Abbildungsverzeichnis	3
1. Einleitung	4
2. Chronologische Zusammenfassung	5
2.1. Jahr 2020	5
2.2. Jahr 2021	6
2.3. Jahr 2022	7
2.4. Jahr 2023	8
3. Technischer Bericht	10
3.1. Einleitende Überlegungen	10
3.2. Operationelle Anforderungen	11
3.3. Anknüpfungspunkte für den Einsatz maschinellen Lernens	11
3.4. Einbettung in den Kontext des Fixed Flight-Path Angle Descent (FPA)	12
3.5. Möglichkeiten der Transition zum Ein-Personen-Cockpit	14
3.6. Operationelles Konzept	14
3.7. Anwendungsfall „Approach Briefing“	15
3.8. Gewinnung von Trainingsdaten	17
3.9. Interne Systemarchitektur des Approach Briefing Demonstrators	18
3.10. Funktionale Schnittstelle: Pilot – Assistenzsystem	18
3.11. Informationsbereitstellung (HMI)	19
3.12. Integration	21
3.13. Evaluierung	23
3.13.1. Ablauf	24
3.13.2. Messmethodik	25
3.13.3. Auswertung	27
4. Veröffentlichungen, Vorträge	30
5. Anhang	32

Abkürzungsverzeichnis

ATC	Air Traffic Control
EASA	European Union Aviation Safety Agency
ENRI	Electronic Navigation Research Institute (Japan)
GUI	Graphical User Interface
IFF	Institut für Flugführung, TU Braunschweig
KI	Künstliche Intelligenz
NOTAM	Notice(s) To Air Missions
PF	Pilot Flying
PM	Pilot Monitoring
SPO	Single Pilot Operation
TOD	Top Of Descent
VCA	Virtueller Crewassistent

Abbildungsverzeichnis

Abbildung 1: Erster konzeptioneller Entwurf des virtuellen Crewassistenten	10
Abbildung 2: Höhenprofile der unterschiedlichen Anfluggeometrien „Step-Approach“ (klassisch), „Continuous Descent Approach“ (CDA) und „Fixed Flight-Path Angle Descent“ (FPA)	13
Abbildung 3: Generierung und Verwendung von Trainingsdaten	17
Abbildung 4: Systemarchitektur und Zusammenhang der einzelnen Komponenten des Technologiedemonstrators	18
Abbildung 5: Erster HMI-Entwurf zum Anwendungsfall „Approach Briefing“	20
Abbildung 6: HMI-Entwurf auf einem Electronic Flight Bag (Quelle des Hintergrundes [1]).....	20
Abbildung 7: HMI-Entwurf auf Display der Flugzeugavionik (Quelle des Hintergrundes [2]).....	21
Abbildung 8: Aufnahme-Bildschirm der EFB-Applikation "Approach Briefing Assistant"	22
Abbildung 9: Anzeige der Auswertung der EFB-Applikation "Approach Briefing Assistant"	22
Abbildung 10: Briefing-Material für die Evaluation des Approach Briefing Assistenten. Kartenmaterial nur zu Anschauungszwecken, nicht zur Navigation zu nutzen.....	25
Abbildung 11: Einordnung des SUS-Ergebnisses auf der in [3] gegebenen Skala. Der Durchschnittswert ist markiert.	27
Abbildung 12: Boxplots der Ergebnisse des "System Usability Scale" (SUS) und dem "Trust of Automated Systems Test" (TOAST) aufgeschlüsselt in die Subgruppen "Performance" und "Understanding"	28
Abbildung 13: Präsentation des Anwendungsfalls "Approach Briefing" auf der ILA Berlin 2024.....	31

1. Einleitung

Der vorliegende Abschlussbericht dokumentiert die Ergebnisse und Erkenntnisse eines mehrjährigen Projekts, das sich der Entwicklung eines innovativen virtuellen Crewassistenten widmete. Das Projekt ViCKI (Virtueller Crewassistent - Kontextspezifische Unterstützung für „Reduced Crew Operations“ mittels Künstlicher Intelligenz) verfolgte das Ziel, Konzepte für eine umfassende digitale Assistenz zu entwickeln, die Piloten bei der Durchführung ihrer Aufgaben unterstützt und die Sicherheit sowie Effizienz im Cockpit erhöht. Beteiligt am Projekt waren insgesamt fünf Partner, darunter DIEHL als Verbundführer, Jeppesen als weiterer Industriepartner und die Forschungseinrichtungen TU Darmstadt, das Deutsche Forschungszentrum für Künstliche Intelligenz (DFKI) und schließlich das Institut für Flugführung der TU Braunschweig.

Der Bericht fasst die wesentlichen Schritte des Entwicklungsprozesses zusammen, von der Konzeption und den ersten Analysen über die Prototypenentwicklung bis hin zur abschließenden Evaluierung und Tests. Darüber hinaus wird auf die Zusammenarbeit der verschiedenen Projektpartner eingegangen, deren interdisziplinärer Ansatz maßgeblich zum Erfolg des Projekts beitrug.

Dieser Abschlussbericht dient sowohl als Dokumentation des Entwicklungsprozesses als auch als Grundlage für die weitere Verwertung und mögliche zukünftige Weiterentwicklungen des virtuellen Crewassistenten. Im Folgenden werden zuerst die einzelnen Phasen des Projekts im chronologischen Verlauf beschrieben, um einen Überblick über den allgemeinen Ablauf des Projekts zu geben. Danach folgt der technische Bericht mit Details zu den konzeptionellen Entscheidungen, der Implementierung und den Ergebnissen.

2. Chronologische Zusammenfassung

2.1. Jahr 2020

Im Jahr 2020 begann das ViCKI-Projekt mit dem Ziel, ein Konzept für einen virtuellen Crewassistenten (VCA) für die Cockpitcrew von Verkehrsflugzeugen zu entwickeln. Der Fokus lag eingangs auf der detaillierten Planung und Konzeptionierung, um eine solide Basis für die spätere Entwicklung und Implementierung eines Technologiedemonstrators zu schaffen.

Diese Anforderungen an ein operationelles Konzept wurden in Zusammenarbeit mit verschiedenen Interessengruppen erarbeitet, darunter Fachleute aus der Luftfahrtindustrie, der Europäischen Agentur für Flugsicherheit (EASA) und potenzielle Endbenutzer. Die erarbeiteten Anforderungen berücksichtigten sowohl funktionale Aspekte wie die spezifischen Funktionen der Software als auch nicht-funktionale Anforderungen wie Benutzerfreundlichkeit, Sicherheit und Konformität mit geltenden Normen.

Ein wesentlicher Bestandteil des Jahres 2020 war die Ausarbeitung von möglichen Anwendungsfällen, in denen der Einsatz von künstlich intelligenten Algorithmen einen Mehrwert bietet, die Sicherheit jedoch nicht negativ beeinflusst. Im Zuge dieser Ausarbeitung wurden in wechselnden Gruppenzusammensetzungen mögliche Kandidaten für Anwendungsfälle diskutiert und bewertet. Schließlich wurde sich vonseiten der TU Braunschweig für das sogenannte „Approach Briefing“ entschieden. Während eines solchen obligatorischen Briefings besprechen die Piloten den bevorstehenden Anflug und gehen auf unterschiedliche Themen zu dessen Durchführung ein. Grundidee eines solchen Briefings ist, dass die Piloten über denselben Kenntnisstand verfügen, um Missverständnisse während der Durchführung des Anfluges zu vermeiden. Weiter wird das Briefing auch dazu genutzt die korrekte Konfiguration der Flugzeugsysteme in Bezug auf den Anflug zu überprüfen.

Innerhalb des beschriebenen Anwendungsfalls sollte eine mögliche Funktionalität eines VCA demonstriert werden: Nach dem Aktivieren des VCA wird durch die Piloten das Approach Briefing gegeben und vom VCA aufgenommen. Nach Abschluss des Briefings analysiert der Assistent das Briefing auf Vollständigkeit und gibt anschließend einen Bericht darüber aus, welche Themen angesprochen und welche nicht adressiert wurden. Mit dieser Unterstützung soll sichergestellt werden, dass kein Thema während des Briefings vergessen wurde. Entsprechend ist eine robuste Spracherkennung für den VCA erforderlich. Zur Demonstration einer zeitnahen Implementierung sollte der VCA als Electronic Flight Bag (EFB) Applikation umgesetzt werden. Detaillierte technische Informationen folgen in Kapitel 3.

Anhand des Anwendungsfalls „Approach Briefing“ wurde Ende 2020 auch mit der Entwicklung eines Konzepts für eine explorative Studie begonnen. Ziel dieser Studie sollte sein, die Verschiedenheit von Approach-Briefings unterschiedlicher Crews zu untersuchen und die Notwendigkeit einer Briefing-Assistenz in komplexen Szenarien zu zeigen. Aufgrund der COVID-19-Pandemie und der damit verbundenen Einschränkungen wurde diese Vorstudie zuerst verschoben. Stattdessen wurden alternative Evaluationsansätze erarbeitet, um den veränderten Bedingungen Rechnung zu tragen. Hierzu gehörten die Anpassung des Zeitplans sowie die Entwicklung neuer Evaluationsmethoden, um sicherzustellen, dass die Evaluationsstudien erfolgreich durchgeführt werden können, sobald die Bedingungen es erlauben würden.

Die Pandemie beeinflusste auch die Arbeitsweise und die Projektkoordination. Die Teamarbeit im Jahr 2020 fand überwiegend in Form von Remote-Arbeit und virtuellen Meetings statt. Diese Umstellung erforderte Anpassungen in der Kommunikation und Koordination innerhalb des Projektteams. Trotz der Herausforderungen konnten innerhalb mehrerer Workshops signifikante Fortschritte in der Definition des VCA erzielt werden.

Zusätzlich wurden erste technische Studien und Machbarkeitsanalysen durchgeführt. Diese Studien dienten der Auswahl geeigneter Technologien und der Vorbereitung der technischen Infrastruktur für die kommenden Phasen des Projekts. Die Untersuchungen umfassten die Evaluierung verschiedener Entwicklungswerkzeuge und -plattformen, um sicherzustellen, dass sie den Anforderungen des Projekts entsprechen und die Grundlage für eine erfolgreiche Systemintegration und -validierung bilden.

Insgesamt war das Jahr 2020 durch Planungs- und Vorbereitungsarbeiten geprägt. Trotz der durch die Pandemie verursachten Einschränkungen konnte das Projektteam wesentliche Fortschritte bei der Definition der Systemanforderungen, der Planung der Evaluationsstrategien und der technischen Vorbereitung erzielen. Diese Vorarbeiten bildeten die Grundlage für die erfolgreiche Umsetzung der Entwicklungsaktivitäten und die Erreichung der Projektziele in den folgenden Jahren.

2.2. Jahr 2021

Im Jahr 2021 setzte das ViCKI-Projekt seine Arbeit mit der eigentlichen Entwicklungsphase fort, nachdem die grundlegenden Planungen und Vorbereitungen im Jahr 2020 abgeschlossen worden waren. Eingangs wurden die Entwicklungsarbeiten koordiniert und gemeinsam eine Systemarchitektur mit den Verantwortlichkeiten der einzelnen Partner entworfen. Insbesondere mit dem DFKI wurden verschiedene Ansätze diskutiert, mit denen eine zufriedenstellende Umsetzung des Demonstrators möglich erschien. Technisch betrachtet gab es im Kontext des Forschungsprojekts einige Limitationen, die zum Ausschluss verschiedener Ansätze führten. Besonders hervorzuheben ist dabei die herausfordernde Beschaffung von passenden Trainingsdaten für die Spracherkennung.

Entsprechend früh wurde durch das Institut für Flugführung mit der Sammlung von Trainingsdaten begonnen. Erforderlich waren möglichst viele Aufnahmen von authentischen Approach Briefings sowie deren Transkriptionen. Beeinträchtigt durch die Pandemie wurden dabei verschiedene Ansätze gewählt: Einerseits wurden Piloten innerhalb von Online-Meetings gebeten zu einem bestimmten Szenario ein Briefing zu halten. Sobald die Umstände es zuließen, wurden Piloten andererseits auch in den Cockpitsimulator des IFF eingeladen, um vor Ort hochwertige Aufnahmen der Briefings anzufertigen, die später auch für das Training verwendet werden konnten. Von den Online-Briefings wurden lediglich die Transkriptionen verwendet, da die Aufnahmequalität nicht den Ansprüchen entsprach.

Weiter wurde auch mit der Konzeptionierung der Benutzerschnittstelle der EFB-Applikation begonnen. Dafür wurden zuerst Abbildungen von möglichen Integrationen einer Applikation sowohl innerhalb eines EFBs als auch als langfristige Perspektive innerhalb der fest verbauten Flugzeugavionik erstellt. Die Gestaltung der Benutzeroberfläche orientierte sich an den Normen der Europäischen Agentur für Flugsicherheit (EASA) und wurde in enger Abstimmung mit Verkehrspiloten optimiert. Dabei wurde besonderer Wert auf die Benutzerfreundlichkeit gelegt, einschließlich der Verbesserung der Sichtbarkeit und Lesbarkeit von Texten und Schaltflächen auf dem Bildschirm.

Weitere Arbeiten in diesem Jahr konzentrierten sich auf die Entwicklung der grundlegenden Softwarekomponenten der EFB-Applikation. Ein bedeutender Fortschritt im Jahr 2021 war die erste Integration der Sprachverarbeitungs-komponenten in das System. Hierzu wurden Technologien zur Spracherkennung und kontextuellen Analyse implementiert. Diese Komponenten wurden in Zusammenarbeit mit dem DFKI entwickelt und ermöglichten es, gesprochene Informationen während des Briefings zu verarbeiten und relevante Daten zu extrahieren. Die Implementierung umfasste die Integration des Spracherkennungssystems „Vocon“ sowie des kontextuellen Analysewerkzeugs „Lodex“. Diese Technologien bildeten die Grundlage für die künftige Kontextualisierung und Verarbeitung der in den Briefings enthaltenen Informationen.

Zusätzlich wurden die ersten Schritte zur Evaluierung des Systems vorbereitet. Dazu gehörten die Definition von Evaluationskriterien und die Entwicklung von Bewertungsinstrumenten, um die Leistungsfähigkeit und Benutzerakzeptanz des Systems zu beurteilen. Die Herausforderungen des Jahres 2021 umfassten weiterhin die Anpassungen an die pandemiebedingten Einschränkungen, die die Teamarbeit und die Durchführung von Tests beeinflussten. Die Projektkoordination wurde weiterhin durch virtuelle Meetings und Remote-Arbeit geprägt, was zusätzliche Anstrengungen zur Sicherstellung der Projektfortschritte erforderte.

Zusammenfassend lässt sich sagen, dass das Jahr 2021 durch den Fortschritt bei der Entwicklung und Integration der Softwarekomponenten sowie durch die Vorbereitung der Evaluationsmaßnahmen geprägt war. Die erfolgreiche Umsetzung der Softwarefunktionen, die Integration der Sprachverarbeitungstechnologien und die Vorbereitung auf die Evaluierung legten die Grundlage für die weitere Entwicklung und Testung des Systems im kommenden Jahr.

2.3. Jahr 2022

Im Jahr 2022 wurde die Entwicklung und Verfeinerung der virtuellen Crewassistenz fortgesetzt, wobei der Schwerpunkt auf der Integration, Verbesserung der Systemkomponenten und der Vorbereitung der Evaluation lag. Dieses Jahr war entscheidend für die Konsolidierung der bisherigen Ergebnisse und die Vorbereitung auf die abschließenden Tests und Evaluierungen.

Zu Beginn des Jahres lag der Fokus auf der Erweiterung und Verfeinerung der Benutzeroberfläche und der Funktionalität der EFB-Applikation. Die während der vorhergehenden Jahre entwickelten Prototypen wurden weiter optimiert, um den Anforderungen der Benutzerfreundlichkeit und der EASA-Normen noch besser gerecht zu werden. Es wurden umfassende Anpassungen vorgenommen, die sowohl die visuelle Gestaltung als auch die ergonomische Bedienbarkeit der Applikation betrafen. Besonderes Augenmerk galt der Verbesserung der Interaktionselemente und der Anpassung der Benutzeroberfläche an die Bedürfnisse der Verkehrspiloten.

Parallel zur Weiterentwicklung der Benutzeroberfläche wurden signifikante Fortschritte bei der Integration und Optimierung der Sprachverarbeitungstechnologien erzielt. Die Sprachverarbeitungssysteme „Vocon“ und „Lodex“ wurden weiter verfeinert, um die Genauigkeit und Relevanz der kontextuellen Analyse zu verbessern. Diese Technologien ermöglichten es, gesprochene Briefings zu erkennen und relevante Informationen zu extrahieren. Ein zentraler Bestandteil der Arbeit war die Anpassung der Algorithmen zur Verbesserung der Kontextualisierung und der Erkennung von Fachtermini in verschiedenen Abschnitten der Briefings.

Im Jahr 2022 wurde auch das Konzept eines virtuellen Dialogpartners vorangetrieben. Die Interaktion zwischen der Approach-Briefing-Assistenz und den Piloten sollte dabei so intuitiv und leichtgewichtig wie möglich erfolgen, um keine signifikante zusätzliche Belastung im operationellen Betrieb darzustellen. Dennoch wurde sich entschieden, eine nicht-sprachgesteuerte Aktivierung der Assistenzfunktion einzuführen, um eine unbeabsichtigte Aktivierung durch Sprachbefehle zu verhindern. Dies trug weiter zur Reduzierung möglicher Akzeptanzprobleme bei, da die Piloten nun durch explizites Drücken von Schaltflächen die Assistenz aktivieren konnten, anstatt dass das System kontinuierlich auf Sprachbefehle lauschte.

Zusätzlich wurden die kontextspezifischen Informationsaufbereitungsalgorithmen weiterentwickelt. Die Zusammenarbeit mit dem DFKI ermöglichte es, die Integration der kontextuellen Vollständigkeitsanalyse zu optimieren. Das System konnte nun besser mit verschiedenen Synonymen und Fachtermini umgehen, was zu einer signifikanten Verbesserung der Analysegenauigkeit führte. Die Arbeit umfasste die Verfeinerung der Algorithmen, um die Trennschärfe bei der Erkennung relevanter Informationen in den Briefings zu erhöhen.

Ein wichtiger Meilenstein im Jahr 2022 war die Fertigstellung der funktionalen und regulatorischen Anforderungen für die EFB-Applikation. Die Definition der Funktionalitäten für die Demonstration der virtuellen Crewassistenz wurde abgeschlossen. Dies ermöglichte es, die entwickelten Komponenten gemäß den festgelegten Anforderungen zu testen und zu bewerten. Das Erreichen des Meilensteins für die funktionale Demonstration stellte einen wesentlichen Schritt zur Vorbereitung auf die abschließende Evaluierung dar.

Die Integration der entwickelten Komponenten und die Vorbereitung auf die abschließende Evaluation waren zentrale Aktivitäten im Jahr 2022. Die Integration der Software- und Hardwarekomponenten wurde erfolgreich abgeschlossen, und die gesamte Verarbeitungskette konnte auf den Systemen des Simulators an der TU Braunschweig in Betrieb genommen werden. Dies ermöglichte erste Tests und die Validierung der Systemleistung im simulierten Betrieb.

Im Rahmen der Vorbereitung auf die Evaluation wurde ein Evaluationsszenario entwickelt und die entsprechenden Bewertungsinstrumente erstellt. Ein Fragebogen zur Bewertung der virtuellen Crewassistenz wurde in Zusammenarbeit mit der TU Darmstadt entwickelt. Dieser Fragebogen beinhaltete sowohl selbst generierte Fragen als auch etablierte Testfragen wie den Trust of Automated Systems Test (TOAST) und die System Usability Scale (SUS).

Im Überblick war das Jahr 2022 durch wesentliche Fortschritte bei der Optimierung der Systemkomponenten, der Verfeinerung der Sprachverarbeitungstechnologien und der Vorbereitung auf die abschließende Evaluierung geprägt. Die erfolgreiche Integration und Tests der entwickelten Komponenten sowie die Erstellung der Evaluationsinstrumente stellten die Weichen für die abschließende Optimierung und Bewertung des virtuellen Crewassistenten im kommenden Jahr.

2.4. Jahr 2023

Im Jahr 2023 konzentrierte sich das ViCKI-Projekt auf die abschließende Verfeinerung und Evaluierung der virtuellen Crewassistenz. Die Arbeiten in diesem Jahr umfassten sowohl die letzte Phase der Systemoptimierung als auch die Evaluierung, um die Leistung und Benutzerfreundlichkeit des Systems zu validieren.

Zu Beginn des Jahres wurden die bereits entwickelten Komponenten intensiv getestet und weiter optimiert. Dies umfasste insbesondere die Weiterentwicklung der Electronic Flight Bag (EFB)-Applikation und der grafischen Benutzeroberfläche. Die im Vorjahr vorgenommenen Anpassungen wurden weiter verfeinert, um sicherzustellen, dass alle visuellen und ergonomischen Anforderungen der EASA-Normen vollständig erfüllt wurden. Die Benutzeroberfläche wurde weiter optimiert, wobei insbesondere der Kontrast der Schrift zum Hintergrund verbessert wurde, um eine bessere Lesbarkeit zu gewährleisten. Die überarbeitete Bedienoberfläche wurde in einem Workshop mit Human-Machine-Interface-Experten der EASA besprochen.

Die Integration der Sprachverarbeitungstechnologien, insbesondere der Spracherkennungssysteme „Vocon“ und der kontextuellen Vollständigkeitserkennung durch „Lodex“, wurde weiter optimiert. Auch wenn das System zu Beginn noch Optimierungsbedarf aufwies, konnte eine stabilere und funktionalere Version des Systems für die geplante Evaluation im Jahr 2024 fertiggestellt werden. Die Evaluationsmethode wurde zusammen mit dem Fachbereich Ingenieurspsychologie der TU Darmstadt entwickelt, und alle notwendigen Unterlagen für die Evaluation wurden bereitgestellt. Eine Vorevaluation mit einem Verkehrspiloten aus dem Umfeld des IFF wurde durchgeführt, um den Evaluationsablauf zu testen und zu validieren.

Im Bereich der kontextspezifischen Informationsaufbereitung wurde die Modellierung abgeschlossen. Das Programm „Lodex“ wurde erfolgreich in das System integriert und für die spezifischen Anforderungen des Approach Briefings optimiert. Die Herausforderung bestand darin, Fachtermini korrekt zuzuordnen, auch wenn sie in verschiedenen Abschnitten des Briefings auftauchten. Durch die Verwendung von sogenannten „Word Embeddings“ konnte die Genauigkeit der Themenzuordnung verbessert werden. Weiter wurde vom IFF-Team eine Synonymliste mit Fachbegriffen beige-steuert, um die Erkennungsrate weiter zu verbessern.

Die Entwicklung des virtuellen Dialogpartners wurde abgeschlossen. Die Interaktion zwischen der Approach-Briefing-Assistenz und den Piloten wurde sowohl software- als auch hardwaretechnisch realisiert.

Die Integration der gesamten Systemarchitektur wurde im Jahr 2023 erfolgreich abgeschlossen. Das Gesamtsystem, bestehend aus der GUI, der Spracherkennung und der kontextuellen Vollständigkeitsanalyse, wurde auf dem Simulator an der TU Braunschweig in Betrieb genommen und getestet. Ein „Design Freeze“ wurde festgelegt, um den Stand der KI-Systemoptimierung bis Dezember 2023 einzufrieren. Weitere Anpassungen wurden ausgeschlossen, um die Evaluationsstudie nicht zu beeinträchtigen.

Die abschließende Evaluierung des virtuellen Crewassistenten begann im Laufe des Jahres. Erste Tests im Simulator der TU Braunschweig wurden durchgeführt, und verschiedene Sprecher wurden gebeten, Texte einzusprechen, um die Systemleistung zu prüfen. Diese Tests wurden erfolgreich abgeschlossen und bereiteten den Weg für die Rekrutierung der ersten Studienteilnehmer zu Beginn des Jahres 2024.

Zusammengefasst war das Jahr 2023 geprägt von der finalen Optimierung der Systemkomponenten, der Durchführung umfangreicher Tests und der Vorbereitung auf die abschließende Evaluierung des virtuellen Crewassistenten. Die erfolgreiche Integration und die Durchführung der ersten Tests stellten sicher, dass das System bereit für die Abschlussstudie im Jahr 2024 war.

Nach dem chronologischen Überblick folgt nun der technische Bericht, der die erwähnten Meilensteine im Detail erläutert und Besonderheiten herausstellt. Auch im technischen Bericht wurde grundsätzlich darauf geachtet, die Neuerungen in zeitlich korrekter Reihenfolge darzustellen.

3. Technischer Bericht

3.1. Einleitende Überlegungen

Zu Beginn des Projektes wurde ein konzeptioneller Entwurf für eine mögliche technische Umsetzung des virtuellen Crewassistenten (VCA) erarbeitet. Dieser wird in Abbildung 1 veranschaulicht und nachfolgend beschrieben.

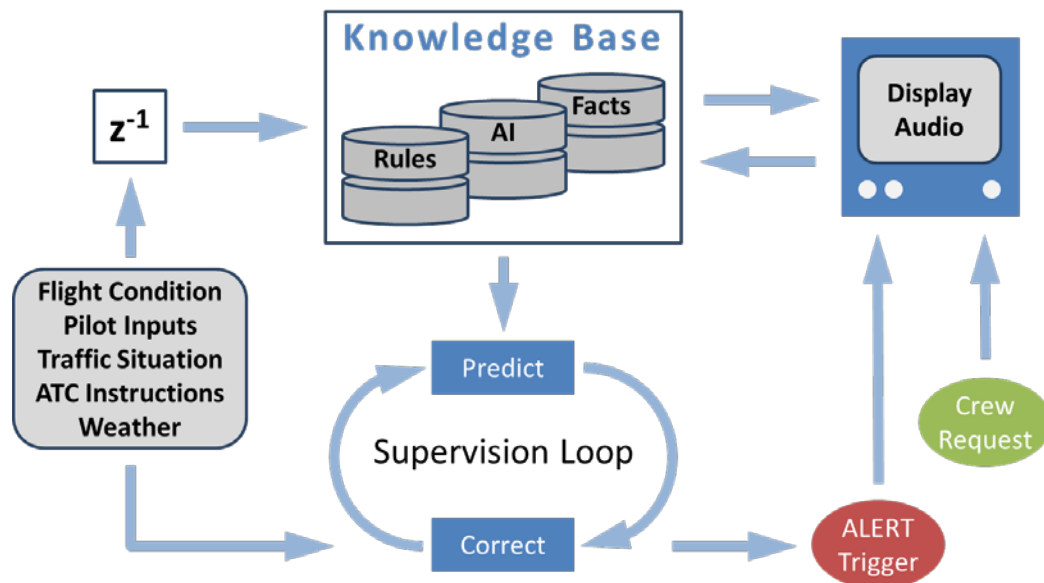


Abbildung 1: Erster konzeptioneller Entwurf des virtuellen Crewassistenten

Ziel des Konzeptes war es, eine überwachende Assistenz und ggf. eine notwendige Alarmierung der Crew während eines Fluges zu realisieren. Aus der Sicht der Piloten soll der VCA als virtueller Dialogpartner zur Verfügung stehen.

Kernelement ist der „Supervision Loop“: Ausgehend von der aktuellen Situation (Flugzustand, Piloteneingaben, Verkehrssituation, etc.) wird ein zukünftiger Zustand prädiziert, der unter normalen Umständen nach einem bestimmten Zeitintervall eintreten müsste. Grundlage für diese Prädiktion ist eine Wissensdatenbank die neben Regeln, Modellen und Fakten auch maschinell erlerntes Wissen beinhaltet.

Nach Ablauf des Zeitintervalls wird überprüft, inwieweit der prädizierte Zustand von dem tatsächlich eingetretenen Zustand abweicht. Laufen Prädiktion und Realität zunehmend auseinander, wird ein Alarm generiert und auf die Wissensdatenbank zurückgegriffen, um die Piloten situationsangepasst und nachvollziehbar zu informieren.

Neben der alarm-getriggerten Informationsausgabe sollen auch die Piloten selbst die Möglichkeit haben, Informationen abzurufen. Diese werden dann entsprechend aufbereitet in visueller Form zur Verfügung gestellt.

Der geschilderte Entwurf ist in einem frühen Projektstadium und beschreibt in erster Linie ein übergeordnetes Leitbild zur Identifikation von Komponenten, Schnittstellen und Algorithmen.

3.2. Operationelle Anforderungen

Um ein tiefergehendes operationelles Konzept zu entwerfen, wurden zuerst die Anforderungen an die Entwicklung und spätere Nutzung eines VCA aus Sicht der operationellen Anwendung identifiziert. Da im Rahmen des Projektes ViCKI aus Kapazitätsgründen nicht alle Phasen eines Fluges berücksichtigt werden konnten, wurden in einem Workshop unter der Leitung des Projektpartners Jeppesen Anwendungsfälle identifiziert, für die das Potenzial eines KI-gestützten VCA genauer untersucht werden sollten.

Eine vielversprechende Anwendung für den VCA ist die Überwachung und gegebenenfalls inhaltliche Ergänzung des Approach Briefings (dt. Anflugbesprechung), das Piloten in einem klassischen Zwei-Personen-Cockpit - in der Regel ohne Zeitdruck - vor dem Anflug auf einen Flughafen durchführen. Die Struktur dieses Briefings ist in weiten Teilen durch den Flugzeughersteller bzw. die Standard Operating Procedures des Betreibers vorgegeben, beinhaltet aber auch Punkte, die situationsangepasst angesprochen werden. Dazu zählen unter anderem besondere Windverhältnisse, Terrain oder Landeklappenstellung. So ist es beispielsweise bei einem Anflug auf den Flughafen Bremen nicht notwendig, sich über Geländehindernisse auszutauschen. Bei einem Anflug auf Innsbruck ist dieser Punkt auf Grund der Alpen dagegen von hoher Bedeutung.

Im Grundsatz ist das Briefing der Piloten deren in Worte gefasstes Situationsbewusstsein. Wird ein Punkt angesprochen, ist dieser den Piloten auch bewusst. Wurde etwas nicht gebrieft, besteht zumindest die Gefahr, dass die Situation in diesem Punkt nicht vollständig erfasst wurde. An dieser Stelle sollte der VCA in der Lage sein, aus einem "großen Erfahrungsschatz" zu schöpfen, was "üblicherweise" in bestimmten Situationen und an bestimmten Orten gebrieft werden müsste und warum. Im Cockpit würden die Piloten den VCA als ein zusätzliches und hocheffizientes Besatzungsmitglied wahrnehmen, welches das Briefing mitverfolgt. Spricht die Besatzung im Zuge des Briefings alle situativ relevanten Punkte an, gibt es keine weiteren Hinweise des VCA. Fehlt etwas, könnte beispielhaft der folgende Hinweis gegeben werden: "Windscherung im Anflug möglich, Fluggeschwindigkeit beachten!". Ein Beispiel hierfür ist der Anflug auf den Flughafen Heraklion (Kreta, Griechenland) mit seiner zum Teil anspruchsvollen Windsituation im Endanflug auf die Landebahn 27.

Um einer möglichen Informationsüberflutung der Piloten entgegenzuwirken, sollte der VCA nur dann auf etwas aufmerksam machen, wenn es in der aktuellen Situation tatsächlich relevant ist. Damit einhergehen sollte eine gewisse Toleranz, wenn das durchgeführte Briefing vielleicht nicht ganz einem idealen Briefing entsprach, aber weitgehend von einem Verständnis der Situation durch die Piloten ausgegangen werden kann.

3.3. Anknüpfungspunkte für den Einsatz maschinellen Lernens

Im Kontext der Briefing Assistenz konnten vielfältige Möglichkeiten für den Einsatz von Künstlicher Intelligenz identifiziert werden, um der grundlegenden Projektausrichtung Rechnung zu tragen.

Zunächst muss der VCA in der Lage sein, das von den Piloten durchgeführte Briefing im Kontext zu verstehen. Über die reine Spracherkennung hinaus muss erkannt werden, welche Punkte im Briefing angesprochen werden, selbst wenn sie von Crew zu Crew unterschiedlich formuliert werden. So könnte die bereits beschriebene Windscherung am Flughafen Heraklion wie folgt gebrieft werden:

„Windscherung im Anflug möglich, Fluggeschwindigkeit beachten!“. Aber auch mit der umgangssprachlichen Formulierung „Der Wind ist hier tückisch. Wenn wir nicht auf die Fahrt aufpassen, haben wir ein Problem.“ wäre dieser Punkt ausreichend besprochen.

Große Bedeutung kommt der Frage zu, wie ein ideales Briefing beim Anflug auf einen bestimmten Flughafen und unter bestimmten Bedingungen aussehen müsste. Eine einfache Datenbankabfrage wäre nicht zielführend und würde – vereinfacht formuliert – alle möglichen Punkte liefern, von denen einige aktuell vielleicht gar nicht relevant sind und die Piloten eher ablenken. Hier ist ein tiefergehendes Situationsverständnis erforderlich, das auf bereits gemachten Erfahrungen basiert und nicht unbedingt analytisch ermittelt werden kann.

Zuletzt muss ViCKI in der Lage sein, das von den Piloten durchgeführte Briefing zu bewerten. Dies geht über das bloße Abhaken von Schlüsselwörtern hinaus, da ein und derselbe Punkt – trotz der hohen sprachlichen Standardisierung für Cockpitbesatzungen – unter Umständen sehr unterschiedlich formuliert oder im Kontext mit anderen Themen genannt werden kann.

3.4. Einbettung in den Kontext des Fixed Flight-Path Angle Descent (FPA)

Die oben dargestellte Briefing-Thematik sollte zudem in die Erforschung neuartiger Anfluggeometrien wie z. B. dem Fixed Flight-Path Angle Descent (FPA) eingebettet werden. Im Zusammenhang mit diesem Anflugverfahren kann bereits auf mehrjährige Forschung des IFF in Kooperation mit dem Electronic Navigation Research Institute (ENRI) in Japan zurückgegriffen werden. Die Durchführung von FPA-Anflügen sowie deren Integration in den übrigen Flugverkehr stellen in der Regel hohe Anforderungen an Piloten und Flugverkehrskontrolle. Es ist zu vermuten, dass bei der Durchführung von FPA-Anflügen auf einen Flugplatz zukünftig die rechen- und datenintensiven Prozesse in Bezug auf die Gesamtverkehrslage eher am Boden durchgeführt werden. Cockpitseitig stehen jedoch Themen wie Energiemanagement, Lärminderung und Treibstoffersparnis im Vordergrund, so dass auch hier ein gewisses Potenzial für die Nutzung eines virtuellen Crewassistenten zu erwarten ist.

Das Hauptaugenmerk der weiteren Arbeit sollte weiter dem Approach Briefing gelten. Die Kombination beider Themen war jedoch wünschenswert, da sie dieselbe Flugphase betreffen und insbesondere bei zukünftigen Simulator-Studien ohne großen Mehraufwand Daten für zwei Forschungsfelder gewonnen werden konnten. Auf diese Weise konnte im Rahmen von ViCKI ein Beitrag zum Ausbau der Zusammenarbeit zwischen ENRI und dem IFF geleistet werden. Im Sinne von gegenseitigen Daten-, Informations- und Erfahrungsaustausch profitieren beide Seiten. Insbesondere in der Anfangsphase des Projekts wurden dahingehend Daten mit dem Simulator des IFF erhoben, die im Sinne der Kooperationen mit den ViCKI Projektpartnern und ENRI verwendet und geteilt wurden. Im weiteren Verlauf des Projektes verschob sich der Fokus jedoch zunehmend in Richtung des Approach Briefings, zumal in Rücksprache mit den Partnern kein nennenswerter Beitrag an der Berechnung von FPAs durch künstlich Intelligente Algorithmen, die sich dem Machine Learning bedienen, identifiziert werden konnte.

Aus Gründen der Vollständigkeit soll das Konzept des FPA Descents hier dennoch kurz dargestellt werden:

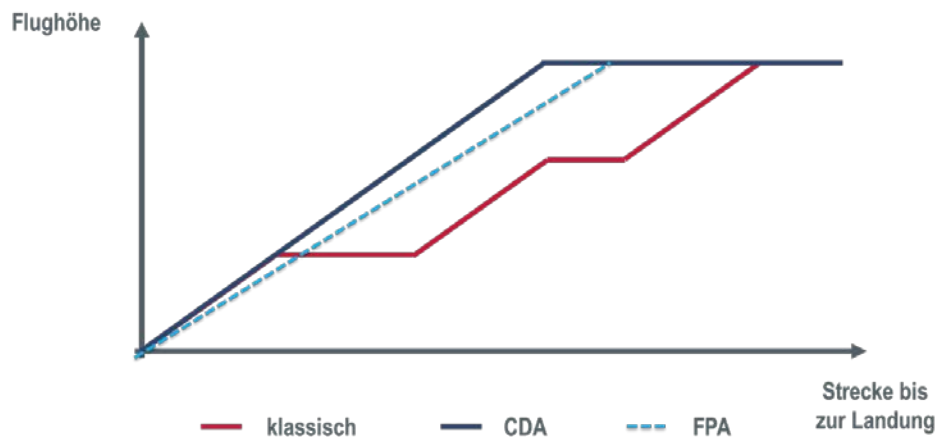


Abbildung 2: Höhenprofile der unterschiedlichen Anfluggeometrien „Step-Approach“ (klassisch), „Continuous Descent Approach“ (CDA) und „Fixed Flight-Path Angle Descent“ (FPA)

Bei heutzutage üblichen Anflugprofilen führt die Flugsicherung das Luftfahrzeug stufenweise von seiner Reiseflughöhe in niedrige Flughöhen (klassisch). Dieses Verfahren wird als Step Descent beschrieben. Der Vorteil des Verfahrens ist die – aus Flugsicherungssicht – gute vertikale Planbarkeit des Sinkflugs. Der Nachteil dieses Verfahrens liegt in den vorhandenen Horizontalflugsegmenten, in denen das Luftfahrzeug mit erhöhter Triebwerksleistung ineffizient in niedrigeren Höhen fliegt.

Demgegenüber steht der Ansatz des Continuous Descent Approaches (CDA). Hierbei wird der Sinkflug erst später begonnen und das Flugzeug sinkt, in vereinfachter Darstellung, kontinuierlich bis zum Erreichen des Endanflugs. Die Treibstoffersparnis wird in diesem Verfahren durch das Verlegen der horizontalen Flugphasen des klassischen Step Approaches auf die Reiseflughöhe erreicht, in der eine höhere Effizienz durch eine geringe Dichte der umgebenden Luft und damit weniger Widerstand gegeben ist. Die Triebwerksleistung verbleibt während dieses Sinkflugs im Leerlauf. Ein entscheidender Nachteil dieses Verfahrens ist die, aus Sicht der Luftverkehrskontrolle, eingeschränkte Planbarkeit der vertikalen Trajektorie, da der vertikale Flugweg und die Fluggeschwindigkeit während des Sinkflugs stark abhängig von der gegenwärtigen Windsituation sind.

Der FPA Descent ist nun ein Ansatz die Vorteile beider Verfahren zu vereinen und die Nachteile zu minimieren. Das Flugzeug hält während des Anfluges einen festen Winkel ein, welcher betragsmäßig jedoch kleiner als der des CDA ist. Dadurch muss die Triebwerksleistung leicht oberhalb der Leerlaufdrehzahl, aber weit unterhalb der Leistung bleiben, die während eines Horizontalflugsegments nötig ist. Auf diese Weise ist dem Flugzeug möglich, einem geometrisch definierten Flugweg zu folgen (Fixed Flight-Path Angle). Der Lotse weiß damit im Vorfeld, wann und in welcher Höhe ein Flugzeug an einem bestimmten Punkt ankommt, was erhebliche Vorteile für die Verkehrsregelung bedeutet.

Da der Sinkflug oberhalb der Leerlaufdrehzahl erfolgt, besteht zudem die Möglichkeit, die Flugzeuge über die Zuweisung einer Geschwindigkeit entlang des Vertikalprofils zu staffeln. Beides ist im Fall der bisher praktizierten CDA nicht möglich. Aus diesem Grund wird CDA in Deutschland nur vereinzelt an wenigen Verkehrsflughäfen und fast ausschließlich bei geringem Verkehrsaufkommen (nachts) eingesetzt. Zusammengefasst verbindet der FPA Treibstoffeffizienz und Lärmvermeidung mit

Planbarkeit des Anfluges und eignet sich somit nach ersten Erkenntnissen als fortschrittliches Anflugkonzept, für zukünftige Ansätze im Bereich der Flugverkehrskontrolle.

3.5. Möglichkeiten der Transition zum Ein-Personen-Cockpit

Der Fokus der Entwicklungen am IFF lag auf den Abläufen in einem konventionellen 2-Personen-Cockpit. Die klassische Rollenverteilung in „Pilot Flying“ und „Pilot Monitoring“ findet sich auch in den Briefing-Prozessen wieder. Im Rahmen des Approach Briefings wird der Pilot Flying den Anflug und die Landung anhand der vorgegebenen Briefing-Struktur durchsprechen, der Pilot Monitoring überwacht das Briefing. Die Integration des VCA ist in diesem Zusammenhang unproblematisch, da er nur dann in den Briefing-Prozess eingreift, wenn etwas falsch oder unvollständig war. Die Verantwortung für die Durchführung eines korrekten Briefings verbleibt letztendlich bei den Piloten.

In einem Ein-Personen-Cockpit ist die Ausgangssituation eine völlig andere. Die klassische Rollenverteilung in „Pilot Flying“ und „Pilot Monitoring“ ist aufgelöst, da die Crew nur aus einem Piloten – dem Pilot Flying oder auch „Mission Manager“ – besteht. Unabhängig davon ist es auch bei Single Pilot Operations (SPO) unabdingbar, dass sich der Pilot der gegenwärtigen Situation bewusst ist und ein mentales Bild für den Anflug aufbaut.

Insofern bei SPO weiterhin ein Briefing vorgesehen ist, gibt es prinzipiell zwei Möglichkeiten, wie dieses Briefing in Verbindung mit einem virtuellen Crewassistenten durchgeführt werden kann. Zum einen kann der Crewassistent den Piloten briefen, der das Briefing dann mitverfolgt und verinnerlicht. Alternativ könnte der Pilot – wie im konventionellen Zwei-Personen-Cockpit – das Briefing durchführen. Der Crewassistent würde dann die Aufgabe des früheren Pilot Monitoring übernehmen, das Briefing überwachen und ggf. korrigieren oder ergänzen. Aus der Sicht des Piloten würde sich in diesem Punkt nur wenig an den konventionellen Abläufen ändern. Der wesentliche Unterschied besteht jedoch darin, dass dem Crewassistenten nun ebenfalls eine hohe „Verantwortung“ für die korrekte Durchführung des Briefings zukommt. Aus diesem Grund bildet die vom IFF verfolgte Strategie, den Crewassistenten zunächst in ein konventionelles Cockpit zu integrieren, einen sinnvollen und für die breite Evaluierung eines Systems wie des VCA notwendigen Zwischenschritt auf dem Weg zu einem zukünftig möglichen Ein-Personen-Cockpit. Der Projektpartner TU Darmstadt, der sich in seinen Entwicklungen auf SPO konzentriert, wurde in diese Überlegungen mit eingebunden und eine entsprechende Empfehlung ausgesprochen.

3.6. Operationelles Konzept

In Zusammenarbeit mit den Projektpartnern ist ein Dokument erarbeitet worden, welches das operationelle Konzept eines VCA zusammenträgt. Aus den eingangs definierten Anforderungen wurden entsprechende konzeptuelle Entscheidungen abgeleitet und verschriftlicht. Die Beiträge der TU Braunschweig beliefen sich insbesondere auf Zuarbeiten im Sinne des Anwendungsfalls Approach Briefing. Aus Gründen der Vertraulichkeit können an dieser Stelle jedoch keine weiteren Details dieses Dokuments veröffentlicht werden.

3.7. Anwendungsfall „Approach Briefing“

Die TU Braunschweig hat die Entwicklung des Anwendungsfalls Approach Briefings initiiert und maßgeblich vorangetrieben. In diesem Unterkapitel soll der Anwendungsfall im Detail vorgestellt werden, da sich die weiter beschriebenen Arbeiten auf eben diesen beziehen.

Ein Approach Briefing stellt im Luftfahrtkontext eine verbale Darstellung des geplanten Sink- und Anfluges dar. Im Rahmen von Multi-Crew-Cockpits ist der Zweck eines solchen Briefings, dass die Piloten ein einheitliches Verständnis des bevorstehenden Flugvorhabens bekommen. Solche Briefings folgen einer vorgegebenen Struktur mit verschiedenen Unterpunkten. Einige dieser Punkte sind verpflichtend anzusprechen, andere lediglich bei Bedarf und soweit sie für das aktuelle Flugvorhaben sinnvoll sind.

Der in ViCKI entwickelte VCA soll das gegebene Approach Briefing akustisch mitverfolgen. Ziel ist es, die Vollständigkeit des Briefings zu ermitteln. Die Richtigkeit der Inhalte wird durch den Assistenten nicht geprüft. Letzteres bleibt in der Verantwortung der Piloten. Das Approach Briefing hat einen definierten Anfang (z. B. durch den Ausruf „Briefing“) und ein definiertes Ende (z. B. mit dem Ausruf „Fragen?“).

Während des Briefings kommt es nur auf das gesprochene Wort an. Mimik und Gestik wirken allenfalls unterstützend, werden aber vom VCA nicht sensorisch verarbeitet.

Die Struktur des Briefings ist fest vorgegeben. Die verschiedenen Teile des Briefings sind an Schlüsselwörtern erkennbar. Was danach angesprochen werden muss bzw. kann, ist ebenfalls standardisiert. Beispielsweise würde das Schlüsselwort „Wetter“ bedingen, dass danach über Wolken, Sicht, Vereisungsgefahr usw. gesprochen wird.

Die Dauer des Briefings liegt bei etwa 3-5 Minuten. Es sollte vollständig sein, sich dabei aber auf die aktuell relevanten Punkte beschränken.

Bei der Überwachung des Approach Briefings wird der VCA im operationellen Betrieb nur dann einschreiten, wenn ein notwendiger Punkt des Briefings nicht besprochen wurde. Ein „zu ausführliches“ Briefing wird hingegen nicht als falsch eingestuft. Aus diesem Grunde wird auch nicht auf richtig und falsch, sondern auf vollständig oder unvollständig geprüft.

Spracherkennung

Im Rahmen des Approach Briefings wird Sprache verwendet, die von heutigen vortrainierten Sprachmodellen in der Regel nicht erkannt wird. Es wurde diskutiert, ob ein bestehendes Sprachmodell um „Cockpitsprache“ ergänzt, oder ob die Cockpitsprache als eine gänzlich neue Sprache aufgefasst und von Grund auf für eine Offline-Spracherkennung modelliert werden sollte. Für den zweiten Ansatz sprachen die folgenden Argumente:

1. Die Cockpitsprache ist in ihrem Umfang stark begrenzt. Das gilt sowohl für das Vokabular als auch für die möglichen Verkettungen von Worten. Nach dem Wort „Clouds“ könnten z. B. die Wörter „broken“ oder „overcast“ folgen, nicht aber der Flugzeugtyp. Es wurde vermutet, dass mit diesem Vorwissen das Problem enormer Mengen an Trainingsdaten umschifft werden könnte.
2. Ein vortrainiertes Sprachmodell wird üblicherweise die Worte erkennen, die im Zusammenhang mit den vorausgehenden und folgenden Worten am wahrscheinlichsten scheinen, selbst wenn die absolute Wahrscheinlichkeit des erkannten Wortes gering ist.

Problematisch wird dieser Umstand, wenn ein bestimmtes Wort der Cockpitsprache gar nicht oder falsch erkannt wird, da dieses im Moment am wahrscheinlichsten schien. Daher wird besonders auf das kontextuelle Verständnis von Sprache trainiert, um Missverständnisse möglichst zu vermeiden.

Auf Grundlage dieser Überlegungen wurde sich für die Entwicklung einer Offline-Spracherkennung mit eigenem Wortschatz entschieden.

Überprüfung auf Vollständigkeit

Die Überprüfung des Approach Briefings auf Vollständigkeit geht über einen reinen Vergleich mit einem Referenz-Briefing hinaus. Die Reihenfolge der Briefing-Items und konkrete Formulierungen können variieren, jedoch sind die übergeordnete Struktur und die Inhalte gleich.

Die Durchführung des Approach Briefings kann als sogenannter schwach strukturierter Workflow aufgefasst und entsprechend modelliert werden. Im Gegensatz zu idealen Prozessketten kann die Reihenfolge bei schwach strukturierten Workflows beliebig sein, jedoch wird die Einhaltung von Abhängigkeiten innerhalb des Workflows überprüft. Die entsprechenden Regeln dazu werden aus Expertenbefragungen abgeleitet.

Realitätsnähe

Für den ersten Entwicklungsschritt eines akustisch arbeitenden Cockpitassistenten musste angenommen werden, dass Briefings stets in englischer Sprache gehalten werden. Die Kenntnis, dass Crews gleicher Muttersprache ihre Briefings häufig in einem Mix aus Muttersprache und Englisch halten ist dahingehend nicht darstellbar, als dass die Verwendung mehrerer Sprachmodelle mit unterschiedlichen Sprachen notwendig wäre, was den umfänglichen Rahmen dieses Projekt überstiegen hätte. Vor dem Hintergrund, dass Englisch die Amtssprache der Luftfahrt ist und Verkehrspiloten über eine Englischqualifikation verfügen, lässt sich diese initiale Vereinfachung jedoch gut im Hinblick auf die Realitätsnähe rechtfertigen.

Bezüglich des Referenzbriefings wurde sich auf das Briefing des Herstellers Airbus geeinigt. Eine Abwandlung bzw. Erweiterung dieser Struktur auf andere Hersteller bzw. besondere Verfahren einzelner Fluggesellschaften ist für die Zukunft vorstellbar, stellt aber keine Änderung des technologischen Grundprinzips des VCA dar.

Aufgrund der hohen Variabilität werden operationelle Unterbrechungen des Briefings (z. B. durch Flugsicherung, Kabinenbesatzung) derzeit nicht abgebildet. Es wurde von einem durchgängigen Briefing ausgegangen

Schnittstellen

Eingabe: Die Spracherkennung wird auf die Erkennung von Schlüsselwörtern aus dem Bereich des Approach Briefings trainiert. Jedes Mal, wenn eines der Schlüsselwörter im Audiosignal detektiert wird, wird dieses als Text an die Vollständigkeitskontrolle weitergeleitet.

Ausgabe: Die Items und Sub-Items der Briefingstruktur sowie wichtige inhaltliche Punkte werden dahingehend überprüft, ob sie im Briefing angesprochen worden sind. Eine Visualisierung dieser Überprüfung erfolgt über ein Benutzeroberfläche (Display).

3.8. Gewinnung von Trainingsdaten

Für das Training der Offline-Spracherkennung ist es aufgrund der von normaler Sprache stark abweichenden Cockpitsprache erforderlich, viele Stunden Sprache in einer realistischen Umgebung (Cockpit) aufzuzeichnen. Um die Datenbasis zu erweitern, ist geplant, zusätzliche Sprachdaten durch Mischung von aufgezeichneter Sprache (am Boden) und einer Original-Geräuschkulisse aus dem Cockpit zu erzeugen. Für den Betrieb der späteren Briefing-Überwachung müssen das gesprochene Approach Briefing der Piloten und ggf. alle Hintergrundgeräusche für die weitere Prozessierung in Echtzeit erfasst werden. Ein portabler Digitalrecorder wurde für diese Aufgaben durch das IFF beschafft. Mit Hilfe eben dieses Rekorders konnten durch das IFF reale Cockpitgeräusche eines Verkehrsflugzeuges aufgezeichnet werden, welche für den o. g. Zweck verwendet wurden.

Insgesamt war ein mehrstufiger Ansatz vorgesehen. Zunächst sollte die Briefing-Assistenz mit Sprache ohne Hintergrundgeräusche erprobt werden. Im nächsten Schritt wurde die Original-Geräuschkulisse hinzugemischt. Zum Entfernen der Hintergrundgeräusche kann das System entweder so trainiert werden, dass es versucht, die verrauschte Sprache zu erkennen, oder zuerst eine Rauschunterdrückung vorgeschaltet wird und die Spracherkennung auf den Ergebnissen dieser operieren kann.

Hinsichtlich der Inhalte der zum Training der Spracherkennung herangezogenen Approach Briefings wurden verschiedene Strategien diskutiert. Zum Training des Sprachmodells wurde ein Referenzbriefing erstellt und von mehreren Sprechern eingesprochen. Diese Sprecher mussten nicht zwingend Piloten sein, ein realistischer Mix der im Cockpit vorkommenden Dialekte/Akzente sollte dennoch abgebildet werden. In einem weiteren Schritt wurden Piloten gebeten, in Abhängigkeit von einer vorher genau geschilderten Situation das Briefing selbst zu formulieren. Diese Daten wurden schließlich für das Training des kontextuellen Erkenners genutzt. Auf Grund der pandemischen Situation war auch hier ein mehrphasiger Ansatz notwendig. Die ersten von Piloten gehaltenen Briefings wurden innerhalb von mehreren Online-Interviews erhoben. Die aufgezeichneten Sprachdaten eigneten sich qualitativ nicht zur Validierung der Spracherkennung, jedoch konnten die Transkriptionen zum Training verwendet werden. Zur Validierung der Spracherkennung wurden in einer zweiten Phase Piloten in den Simulator am IFF eingeladen und gebeten ein Approach Briefing nach der vorgegebenen Situation einzusprechen. Diese Sprachdaten waren in der Aufnahmequalität schließlich ausreichend, um die mit Transkriptionen trainierte Spracherkennung zu validieren.

Die Abbildung 3: Generierung und Verwendung von Trainingsdaten gibt einen Überblick über die erhobenen Trainingsdaten und deren Verwendung.

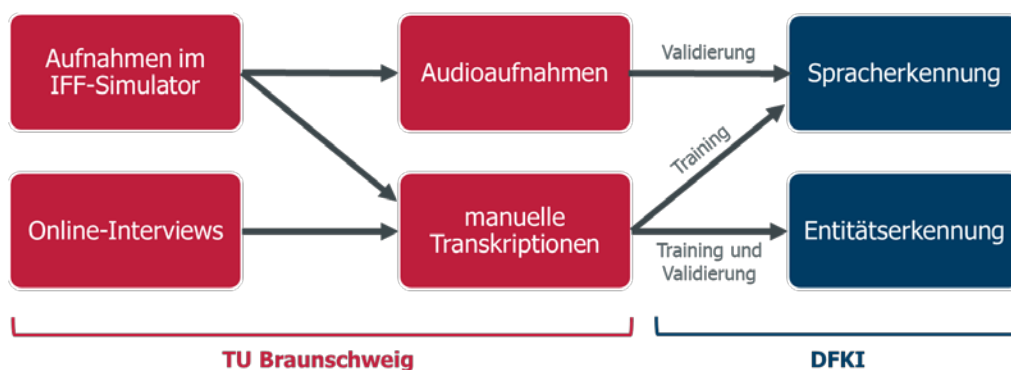


Abbildung 3: Generierung und Verwendung von Trainingsdaten

3.9. Interne Systemarchitektur des Approach Briefing Demonstrators

In Zusammenarbeit mit dem DFKI und der TUDA wurde die Definition der Systemarchitektur für den Anwendungsfall „Approach Briefing“ erarbeitet. Abbildung 4 zeigt den Zusammenhang der einzelnen Komponenten. Für die Transkription und die kontextuale Erkennung wurden die durch das DFKI bereitgestellten Anwendungen „Vocon“ und „Lodex“ integriert. Ein Backend-Server steht dabei für die zentrale Orchestrierung des gesamten Verarbeitungsprozesses bereit. Dieser Server und auch die Android-Frontend-Applikation wurden von der TU Braunschweig entwickelt und integriert.

Die gegebene Grafik veranschaulicht den Datenverarbeitungsprozess: Eingangs wird durch die Frontend-Applikation eine Aufnahme eines Approach Briefings erstellt. Die entsprechende Sprachaufzeichnung wird über den Backend-Server an die Spracherkennungssoftware Vocon weitergeleitet. Nach automatischer Transkription des Textes wird das entsprechende Transkript erneut über den Backend-Server an die kontextuelle Erkennungssoftware Lodex geschickt. Diese extrahiert die angesprochenen Themen des Approach Briefings aus dem Transkript und gibt eine Liste eben dieser an den Backend-Server zurück. Der Backend-Server leitet diese Daten daraufhin an die Frontend-Applikation weiter, die die Ergebnisse dann visuell dem Nutzer präsentiert.

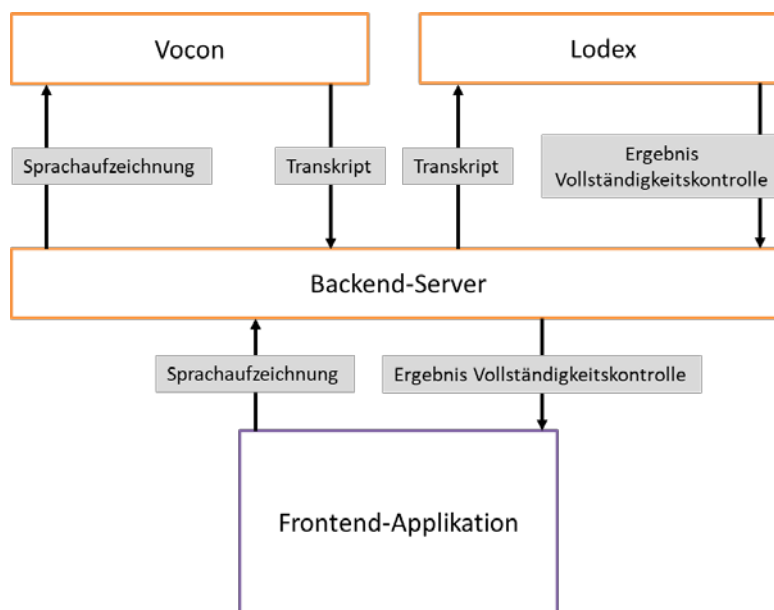


Abbildung 4: Systemarchitektur und Zusammenhang der einzelnen Komponenten des Technologiedemonstrators

3.10. Funktionale Schnittstelle: Pilot – Assistenzsystem

Der VCA tritt als virtueller Dialogpartner mit den Piloten in Kontakt. Dafür wurde eine entsprechende Interaktionsfolge in Bezug auf den Anwendungsfall „Approach Briefing“ definiert:

1. Die Crew entscheidet sich, das Approach-Briefing durch die Assistenzfunktion auf Vollständigkeit überprüfen zu lassen und ruft die entsprechende Assistenzfunktion auf dem EFB (Electronic Flight Bag) auf.
2. Über das HMI auf dem EFB wird dazu die akustische Aufzeichnung des gesprochenen Briefings aktiviert und dies über ein entsprechendes Symbol im HMI angezeigt.

3. Die Crew führt das Briefing wie gewohnt durch. Durch den Assistenten erfolgt während der Durchführung des Briefings keinerlei Rückmeldung hinsichtlich der im Briefing bereits adressierten Briefing-Teile. Auf diese Weise soll sichergestellt werden, dass die Piloten in der Ausführung des Briefings nicht durch die Assistenzfunktion beeinflusst werden.
4. Wenn die Crew das Briefing durchgeführt und alle eventuell offenen Fragen geklärt worden sind, wird die Aufzeichnung des Briefings über das HMI beendet.
5. Nach Beendigung der Aufzeichnung wird das Briefing hinsichtlich seiner Vollständigkeit überprüft und die Ergebnisse auf dem EFB angezeigt.
6. Werden einzelne Teile des Briefings als unvollständig dargestellt, hat die Crew die Möglichkeit, die entsprechend eingefärbten Teile per Touch-Eingabe weiter „auszuklappen“. Beispiel: Der Briefing-Teil „Landing“ wird als unvollständig gebrieft dargestellt. Durch Touch-Eingabe auf das entsprechend-farbig dargestellte Wort „Landing“, werden die Unterpunkte Runway length, width, condition, autobrake setting, etc. sichtbar. In diesem Fall könnte z. B. das entsprechend eingefärbte Wort „Runway Condition“ darauf hinweisen, dass darüber im Briefing nicht gesprochen wurde. Durch die hier beschriebene Möglichkeit wird der Forderung nach AI-Explainability Rechnung getragen.
7. Die Crew kann die entsprechenden Teile nachbriefen und über das HMI entsprechend markieren.
8. Wenn das Briefing nunmehr vollständig abgeschlossen ist, wird die Assistenzfunktion auf dem EFB beendet.

Wichtig ist in diesem Zusammenhang, dass durch die Nutzung des Approach-Briefing-Assistenten nur unwesentlich in die operationellen Abläufe im Cockpit eingegriffen wird.

3.11. Informationsbereitstellung (HMI)

Die TU Braunschweig hat für die Darstellung der Ergebnisse der Vollständigkeitsüberprüfung ein Konzept für eine Mensch-Maschine-Schnittstelle (Human-Machine-Interface, HMI) entwickelt. Auf ein Feedback des Systems durch Audiosignale sollte in Rücksprache mit den Experten des operationellen Betriebs verzichtet werden. Im Cockpit gibt es bereits eine Vielzahl von akustischen Hinweisen, die bei weiterer Ergänzung zu einer Informationsüberfrachtung der Piloten führen könnte. Daher wurde sich im Verbund auf ein Feedback des Systems in visueller Form verständigt.

Abbildung 5 zeigt die angedachte Benutzeroberfläche. Zu sehen sind die einzelnen Überschriften aus der Airbus FCTM-Briefingstruktur in verschiedener Färbung. Diese symbolisiert, ob ein Themenbereich bereits angesprochen wurde. Jede Überschrift hat darüber hinaus auch noch weitere Unterpunkte, die über eine Anwahl der entsprechenden Überschrift ausgeklappt werden kann. Ist ein Themenbereich vollständig besprochen worden, so färbt sich die entsprechende Überschrift von weiß zu grün. Bei teilweiser Vollständigkeit eines Themenbereichs wird die Überschrift gelb gefärbt.

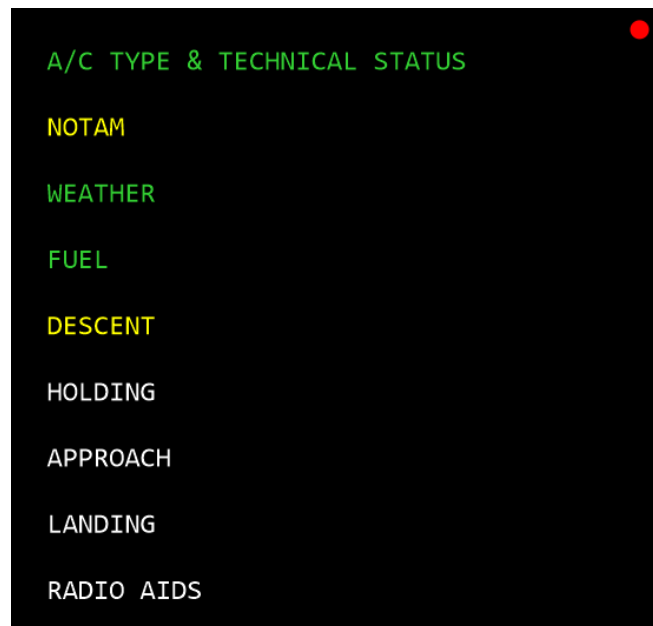


Abbildung 5: Erster HMI-Entwurf zum Anwendungsfall „Approach Briefing“

Der hier gezeigte erste Entwurf beinhaltet noch einen roten Punkt in der rechten oberen Ecke, der als Indikator für eine laufende Aufzeichnung angedacht war. In dieser Projektphase wurde noch davon ausgegangen, dass die einzelnen Themen in Echtzeit während des Briefings farblich entsprechend eingefärbt werden. Aufgrund mehrerer Rücksprachen mit Projektpartnern, Experten aus der Praxis und der EASA wurde jedoch von diesem Ansatz Abstand genommen. Der Hauptgrund für die Änderung von einem „Live-Betrieb“ hin zu einer nachgestellten Verarbeitung lag in der Annahme, dass so das System weniger Ablenkung während des Briefings erzeugt. Mit der Prämisse so wenig wie nur nötig in die operationellen Abläufe im Cockpit einzugreifen, wurde sich schließlich für eine nachlaufende Verarbeitung entschieden.



Abbildung 6: HMI-Entwurf auf einem Electronic Flight Bag (Quelle des Hintergrundes [1])

[1] https://airbus-h.assetsadobe2.com/is/image/content/dam/channel-specific/websites/services/flight-operations/system-upgrades/electronic-flight-bag/A350_Mounting_device.jpg?wid=1920&fit=fit,1&qI=85,0



Abbildung 7: HMI-Entwurf auf Display der Flugzeugavionik (Quelle des Hintergrundes [2])

[2] <https://preview.redd.it/pgu12tz6fll31.jpg?auto=webp&s=739b055419186200bbd4c2c374a1075ecbf4bf41>

Die vorstehenden Abbildungen zeigen exemplarisch die Einbindung des entworfenen HMI-Konzepts mit unterschiedlichen Zeithorizonten. Mittelfristig ist eine Einbindung des VCA als EFB-Applikation denkbar, während die langfristige Perspektive eine Integration innerhalb der fest verbauten Flugzeugavionik zeigt.

3.12. Integration

Die gezeigten Konzepte wurden im Verlauf des Projekts weiter verfeinert und für die Evaluation mithilfe einer EFB-Applikation zur Anzeige gebracht. Basierend auf der Entscheidung für eine EFB Type A Software-Applikation wurden die diesbezüglichen regulatorischen Vorgaben der EASA angewandt, um die visuelle Mensch-Maschine-Schnittstelle der geplanten Briefing-Assistenz weiter auszugestalten. Dabei wurden Aspekte wie Schriftart, -größe, Farbgebung, Anordnung der einzelnen Anzeigeelemente und weitere relevante Faktoren berücksichtigt. Das Ziel bestand darin, sämtliche Designelemente möglichst im Einklang mit den genannten Vorgaben zu entwickeln. Dadurch soll einerseits ein potenzieller Zulassungsprozess berücksichtigt und andererseits sichergestellt werden, dass Probanden, die den Assistenten im späteren Verlauf evaluieren, sich in einer gewohnten Umgebung wieder- bzw. zurechtfinden. Auch in dieser Entwicklungsphase wurden Empfehlungen von Domänenexperten zur Gestaltung des HMI eingeholt und berücksichtigt.

Abbildung 8 zeigt die initiale Ansicht der entwickelten EFB-Applikation. Aus Gründen der Übersichtlichkeit wurde die Anzahl der angezeigten Elemente stark beschränkt sowie die Lesbarkeit durch entsprechend große Schrift adressiert. Weiter sind die relevanten Schaltflächen relativ groß gehalten, um auch im Flug ein sicheres und einfaches Bedienen der Applikation sicher zu stellen.

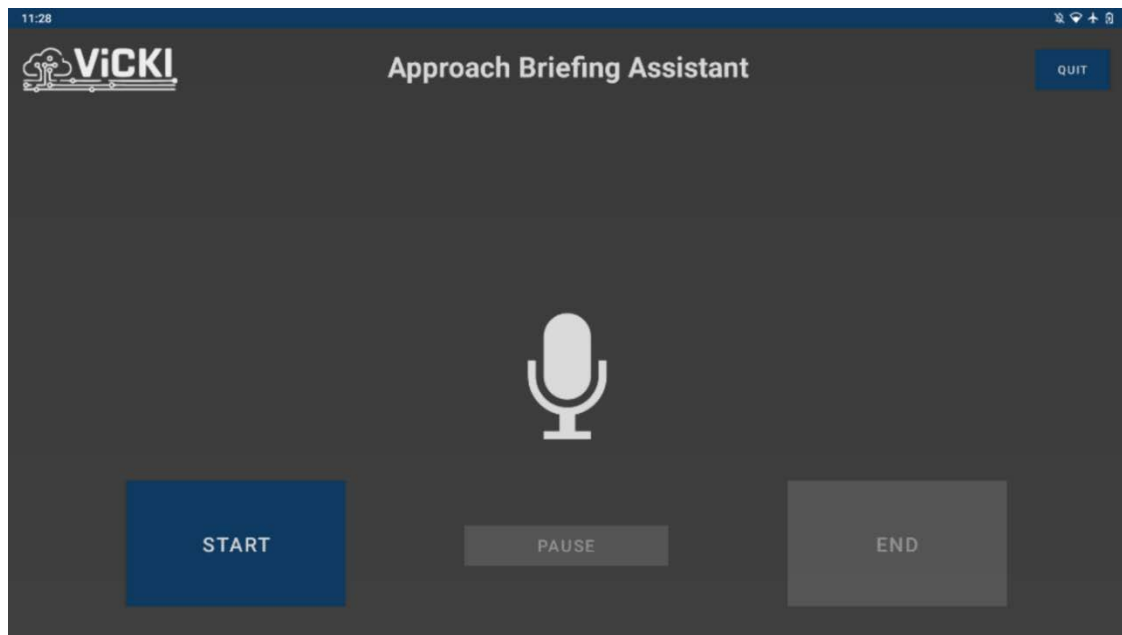


Abbildung 8: Aufnahme-Bildschirm der EFB-Applikation "Approach Briefing Assistant"

Nach einem Druck auf „START“ beginnt die Applikation mit der Aufzeichnung des Briefings. Während der Aufzeichnung pulsiert das Mikrofon-Symbol und indiziert somit eine fortlaufende Funktion des Programms. Es wurde explizit auf die Verwendung von vorbelegten Farben (wie beispielsweise rot) verzichtet, um etwaige Verwechslungen mit aktionsbedürftigen Anzeigen zu vermeiden. Mit Betätigung der Schaltfläche „END“ wird schließlich die Aufzeichnung des Briefings beendet und die entsprechend erstellte Datei an den Backend-Server versandt.

Nach Abschluss der Verarbeitungskette (siehe Kapitel 3.8) zeigt die Applikation die Ergebnisübersicht an, die in Abbildung 9 beispielhaft gegeben ist.

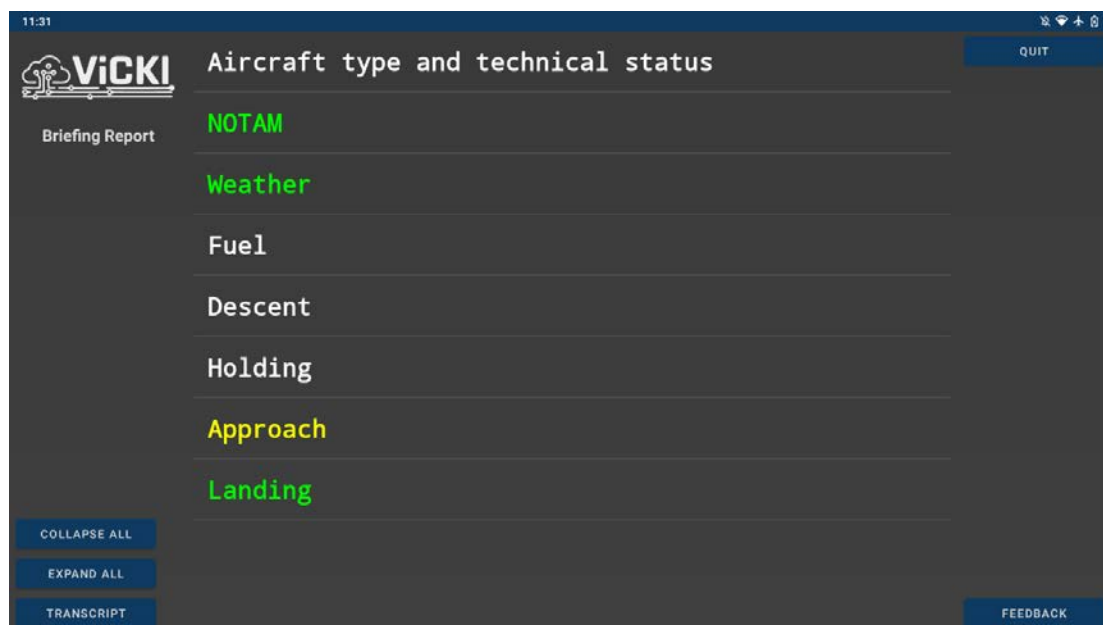


Abbildung 9: Anzeige der Auswertung der EFB-Applikation "Approach Briefing Assistant"

Die vorher beschriebenen Elemente der einzelnen Themen eines Approach Briefings sind auch hier wiederzuerkennen. In diesem Beispiel ist das gegebene Briefing nicht vollständig gewesen. Lediglich die Themen „NOTAM“, „Weather“ und „Landing“ wurden vollständig adressiert und erkannt.

Die Integration der entwickelten Komponenten konnte im Jahr 2023 abgeschlossen werden, sodass eine Evaluierung der entwickelten Systeme möglich war. Die gesamte Verarbeitungskette wurde dabei auf den Systemen des Simulators an der TU Braunschweig in Betrieb genommen und konnte erfolgreich getestet werden. Ferner wurde sich für einen „Design Freeze“ der Entwicklung entschieden, um Effekte einer möglichen Softwareanpassung auf die Evaluationsstudie auszuschließen. Der Stand der Optimierung der KI-Systeme ist somit auf Dezember 2023 festgelegt worden. Weitere Optimierungen seitens des DFKI flossen nicht mehr in die Evaluation mit ein.

3.13. Evaluierung

Die Vorarbeiten aus der Projektanfangszeit für die damals angedachte, aber aufgrund der COVID-19-Pandemie nicht durchführbare Vorstudie konnten größtenteils für die Vorbereitung der Abschlussevaluation wiederverwendet werden. Das Szenario war somit bereits umfänglich definiert und erwies sich nach wie vor als geeignet. Weiter konnte die Evaluation des Anwendungsfalls der TU Darmstadt unabhängig von dem Anwendungsfall „Approach Briefing“ durchgeführt werden, weshalb eine etwaige Synchronisation obsolet geworden war.

Die an der TU Braunschweig vorgesehene Evaluation deckt auch die Arbeitsergebnisse des DFKI ab, da die Anwendungsprogramme dieses Partners Teil der Gesamtapplikation sind.

Der entwickelte virtuelle Crewassistent wurde im Rahmen der abschließenden Evaluationsstudie auf diverse Parameter hin untersucht. Dazu zählen:

- Vollständigkeit (im Rahmen der Evaluationsszenarien),
- Zuverlässigkeit / Wiederholbarkeit,
- Transparenz und Nachvollziehbarkeit,
- Richtigkeit,
- Leistungsfähigkeit / Grenzen der KI-Algorithmen

In Kooperation mit dem Fachbereich Arbeits- und Ingenieurpsychologie an der TU Darmstadt entstand ein Fragebogen, der zum Teil aus selbst generierten Fragen, zum anderen Teil aus etablierten und validierten Testfragen besteht. Insbesondere wurde der Trust of Automated Systems Test (TOAST) und die System Usability Scale (SUS) in die Evaluation mit aufgenommen. Es konnte somit ein Probandenfragebogen erarbeitet werden, der den Studienteilnehmern eine Bewertung des entwickelten VCA im Sinne der oben genannten Parameter ermöglicht.

Die Studie zum Einsatz eines VCA im Bereich der Briefing-Assistenz erfolgte im institutseigenen Simulator, welcher sich in seinem Systemaufbau stark am Cockpit eines Airbus A320 orientiert. Alle wichtigen Systeme sind dabei so detailgetreu nachgebaut, dass für Probanden keine nennenswerten Unterschiede zwischen dem Original und dem Simulator entstehen.

Als Probanden wurden Piloten mit einer kommerziellen Lizenz eingesetzt, einige wiesen auch ein entsprechendes „Type Rating“ (Typenberechtigung) für das Luftfahrzeug Airbus A32F mit Erfahrung im Linienflugbetrieb auf.

3.13.1. Ablauf

Für eine Vergleichbarkeit der Ergebnisse eines jeden Probanden ist ein standardisierter Ablauf einer einzelnen Studiendurchführung von hoher Bedeutung. Daher wurde während der Konzeption der Studie der Ablauf im Detail geplant und mit nicht am Projekt beteiligten Personen aus dem Institut getestet. Ein immer gleicher Ablauf ist deshalb relevant, damit alle Studienteilnehmer auf einen möglichst homogenen Wissenstand gebracht werden, bevor die eigentliche Durchführung der Tests beginnt. Tabelle 1: Ablauf einer Studiendurchführung gibt eine Übersicht über den typischen zeitlichen Verlauf einer Durchführung. Die angegebenen Zeiten wurden im Vorfeld als Richtwerte festgelegt, haben sich aber insgesamt als zutreffende Zeitschätzungen herausgestellt.

Tabelle 1: Ablauf einer Studiendurchführung

Zeitpunkt	Aktion
00:00	Ankommen & Begrüßen
00:10	Stammdatenblatt und Einverständniserklärung
00:15	Einführung in das Projekt ViCKI
00:20	Erklärung und Training EFB-Applikation
00:25	Einführung in das Szenario
00:40	Vorbereitungszeit für das eigene Briefing
00:45	Trainingsdurchlauf
00:50	1. Durchlauf (Briefing nach Vorgabe)
00:55	2. Durchlauf (Briefing frei wie im Berufsalltag)
01:00	Probandenfragebogen
01:10	Nachbesprechung / Interview
01:30	Verabschiedung

Wie der Tabelle zu entnehmen ist, wurde die Gesamtdauer einer Durchführung mit 1:30h angenommen. Nach der Begrüßung wurden die Teilnehmer gebeten, das Stammdatenblatt und die Einverständniserklärung für die Nutzung ihrer Daten zu unterschreiben. Im Anschluss daran wurde eine Einführungspräsentation gezeigt, die grundlegende Informationen zum Projekt ViCKI und den beteiligten Partnern enthielt. Weiter wurde in dieser Präsentation auch auf die entwickelte EFB-Applikation, dessen Bedienbarkeit und das konstruierte Szenario eingegangen. Nach der Präsentation wurde das in Abbildung 10 gezeigte Briefing-Material ausgehändigt. Darin ist zu erkennen, dass die allgemeinen Informationen zu den entsprechenden Themen des Approach Briefings auf der linken Seite in deutscher Sprache aufgelistet sind.

Auf der rechten Seite ist eine der Situation entsprechende Anflugkarte gegeben, die für weitere Details des Briefings erforderlich ist. Die Informationen wurden in deutscher Sprache gegeben, um eine möglichst hohe Varianz in den gesprochenen Briefings zu erreichen. Die Probanden waren so angehalten, die Informationen mit ihren eigenen Worten zu übersetzen und wiederzugeben. Entsprechend standen den Teilnehmern vor dem Versuch alle erforderlichen operationellen Daten für die Flugdurchführung zur Verfügung. Mit diesen Daten konnten die Piloten anschließend ihr eigenes Briefing vorbereiten.

Technical Status	
- Zur Landung liegen keine technischen Einschränkungen vor	
NOTAM	
- RIED DVOR/DME RID 112.2MHZ u/s	
Weather	
- METAR EDDF 010950Z 22015KT 180V240 9999 BKN025 11/07 Q1000 NOSIG=	
Fuel	
- Der Min. Diversion Fuel beträgt gem. OFP 2,3 t. Vor Abflug wurde aus operationellen Gründen entschieden, zusätzliche 0,7 t Kraftstoff mitzunehmen. Dies entspricht dem Verbrauch für einen G/A unter Radar Vectors.	
Descent	
- Es ist von einem ILS- Anflug auf die 25L in EDDF auszugehen - Die momentane Lfz.-Position ist ca. 100 NM südl. des FAP in FL300 - Erwartetes Routing: Es ist von Vectors-to-Final auszugehen - Terrain: Als höchste Grid entlang der restlichen Route konnte 5.600 ft und als höchste MSA 3.500 ft für den betreffenden Sektor ausgemacht werden.	
Holding	
- Die gegenwärtige Verkehrsdichte ist gering. Momentan befinden sich nicht viele Flugzeuge im Anflug. Keines der Flugzeuge musste ein Holding-Verfahren fliegen.	
Approach	
- Type of Approach - Final descent point (identification & altitude) - Glide path - Minima - Missed approach procedure - Alternate considerations	
Landing	
- Runway length, condition and expected wind - LDA vs. RLD (-> 4000 m vs. 1800 m) - Use of autobrake (-> LOW) - Use of reverses (-> IDLE) - Expected taxi route (-> Verlassen der Bahn nach rechts über M17)	

Während des Briefings

- Überschriften nennen
- Englisch
- Eigene Worte
- Kurz & prägnant

	25L	Cat 3b DME	Cat 2 DME	Cat 1 DME	Cat 1 DME	Cat 1 DME	Circling
C	ft - m/km	0 - 75R 95 NA	100 - 300R 95 NA	200 - 400 570	230 - 500 590	200 - 550 570	Not published
D	ft - m/km	0 - 75R 95 NA	100 - 300R 95 NA	200 - 400 570	230 - 500 590	200 - 550 570	Not published

Abbildung 10: Briefing-Material für die Evaluation des Approach Briefing Assistenten. Kartenmaterial nur zu Anschauungszwecken, nicht zur Navigation zu nutzen.

Mit dem darauffolgenden Trainingsdurchlauf wurden die Probanden mit der Bedienung des Approach Briefing Assistenten vertraut gemacht. Daraufhin startete der erste Versuchsdurchlauf, in dem ein Briefing nach der vorgegebenen Briefing-Struktur durchgeführt wurde. Der Versuch selbst startete im Reiseflug ungefähr zwei Flugminuten vor Beginn des Sinkfluges. Die Simulation befand sich dabei im sogenannten Position-Freeze-Modus. Das bedeutet, das Flugzeug war bereits um alle Achsen steuerbar, allerdings behielt es in diesem Modus seine gegenwärtige Position. In dieser Ausgangslage sollten die Piloten ihr Approach Briefing beginnen. Durch den Position-Freeze-Modus liegt kein zeitlicher Druck auf den Probanden, sodass ein zeitunkritisches Briefing erfolgen kann.

Weiter wurde ein zweiter Durchlauf angesetzt, in dem die Piloten ein freies Briefing nach ihrer Gewohnheit durchführen konnten. Dieser Durchlauf diente dem Test der Robustheit der Algorithmen in Bezug auf Testdaten, die signifikant von den Trainingsdaten variierten.

3.13.2. Messmethodik

In der abschließenden Evaluationsstudie kamen zwei wissenschaftlich validierte Bewertungsinstrumente zum Einsatz: der „System Usability Scale“ (SUS) zur Messung der Benutzerfreundlichkeit und der „Trust in Automated Systems Test“ (TOAST), welcher Nutzervertrauen in zwei Dimensionen misst: TOAST_P für Performance (Vertrauen in die Leistungsfähigkeit des Systems) und TOAST_U für Understanding (Verständnis des Systems). Der gesamte Fragebogen zur Evaluation ist diesem Dokument angehängt.

Der SUS ist ein standardisierter Fragebogen zur schnellen und effizienten Bewertung der Benutzerfreundlichkeit eines Systems. Er wurde 1996 von John Brooke entwickelt und ist seitdem ein weitverbreitetes Werkzeug, um die „Usability“ von Systemen zu messen. Der SUS besteht aus zehn Aussagen, die von den Benutzern auf einer Skala von 1 (stimme überhaupt nicht zu) bis

5 (stimme voll und ganz zu) bewertet werden. Diese Aussagen decken verschiedene Aspekte der Benutzererfahrung ab, darunter die Einfachheit der Bedienung, das Vertrauen in die Nutzung des Systems und die allgemeine Zufriedenheit. Um Verzerrungen zu minimieren, wechseln sich positive und negative Aussagen ab.

Die Berechnung des SUS-Scores erfolgt, indem für die ungeraden (positiven) Fragen die Punktzahl um 1 reduziert wird, während bei den geraden (negativen) Fragen die Differenz zwischen 5 und der Bewertung gebildet wird. Die so ermittelten Werte werden summiert und anschließend mit 2,5 multipliziert, um einen Wert zwischen 0 und 100 zu erhalten.

Ein SUS-Score von 68 gilt als durchschnittlich. Werte über 68 weisen auf eine überdurchschnittliche Benutzerfreundlichkeit hin, während Scores darunter Verbesserungspotenzial signalisieren. Der SUS ist ein einfaches, aber effektives Instrument, das aufgrund seiner Einfachheit und statistischen Zuverlässigkeit geschätzt wird und sich zur Bewertung verschiedenster Systeme eignet.

Neben der Benutzerfreundlichkeit sollte auch das Vertrauen in das System bewertet werden. Dazu wurde der TOAST verwendet. Es handelt sich dabei um ein Bewertungsinstrument, das darauf abzielt, das Vertrauen in die Fähigkeit eines Systems, Aufgaben zuverlässig und nachvollziehbar zu erfüllen, zu bewerten. Der TOAST untersucht dabei zwei wesentliche Dimensionen: „Performance“ und „Understanding“.

Die Performance-Dimension misst, inwieweit die Benutzer darauf vertrauen, dass das System seine Funktionen korrekt und effizient ausführt. Es geht dabei um die wahrgenommene Zuverlässigkeit und Effektivität der Systemleistung. Die Understanding-Dimension bewertet hingegen das Vertrauen der Nutzer in ihr eigenes Verständnis des Systems. Hier wird ermittelt, ob die Benutzer das System und dessen Funktionsweise ausreichend nachvollziehen können und wie gut sie in der Lage sind, mit dessen Entscheidungen und Handlungen umzugehen.

Die Ergebnisse des TOAST helfen dabei, ein besseres Verständnis für das Vertrauen der Nutzer in verschiedene Aspekte eines automatisierten Systems zu gewinnen. Diese Erkenntnisse sind besonders wichtig, wenn es darum geht, die Akzeptanz und den erfolgreichen Einsatz solcher Systeme in sicherheitskritischen Bereichen zu verbessern. Der TOAST liefert wertvolle Informationen darüber, ob ein System so gestaltet ist, dass es das Vertrauen der Benutzer stärkt und somit einen reibungslosen und sicheren Betrieb ermöglicht.

Nachfolgend werden die Ergebnisse der 11 Studienteilnehmer dargestellt.

3.13.3. Auswertung

Tabelle 2 zeigt die nach Teilnehmern aufgeschlüsselten Auswertungsergebnisse sowie eine Gesamtauswertung der Skalen SUS, TOAST_P und TOAST_U.

Tabelle 2: Übersicht der Einzelauswertungen der Fragebögen

Teilnehmer	SUS	TOAST_P	TOAST_U
1	97,50	5,60	6,50
2	92,50	5,00	6,00
3	72,50	5,40	5,50
4	82,50	5,20	5,50
5	62,50	3,80	5,75
6	75,00	5,20	5,75
7	87,50	4,80	6,00
8	67,50	4,00	4,75
9	72,50	5,00	6,00
10	87,50	5,20	5,50
11	72,50	4,60	5,75
Minimum	62,50	3,80	4,75
Maximum	97,50	5,60	6,50
Durchschnitt	79,09	4,89	5,73
Median	75,00	5,00	5,75
Standardabweichung	10,57	0,54	0,42

Die **SUS-Werte** reichen von **62,50** bis **97,50**, mit einem Durchschnitt von **79,09**. Dies deutet darauf hin, dass die Mehrheit der Teilnehmer das System als benutzerfreundlich empfand. In der Literatur wird ein solches Ergebnis als „gut“ bewertet, wobei die Nutzbarkeit insgesamt mit dem höchsten Prädikat „akzeptabel“ eingeordnet wird. Der Median liegt bei **75** und die Standardabweichung beträgt **10,57**.

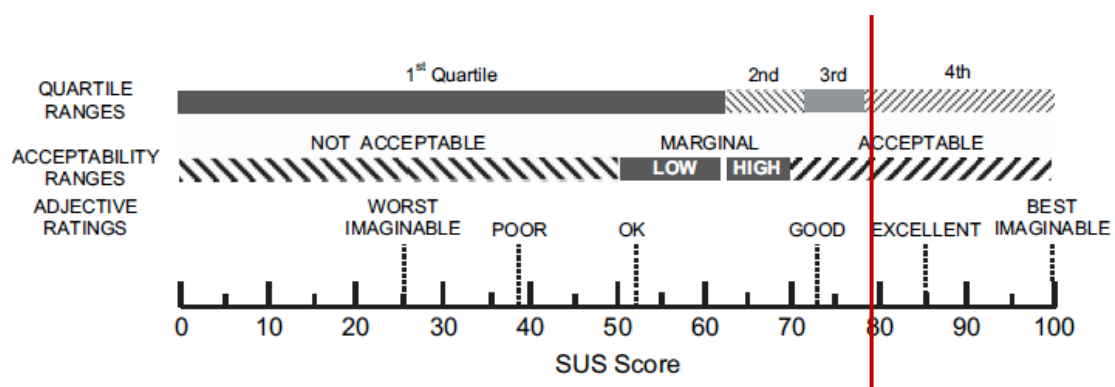


Abbildung 11: Einordnung des SUS-Ergebnisses auf der in [3] gegebenen Skala. Der Durchschnittswert ist markiert.

[3] Bangor, A.; Kortum, P. T.; Miller, J. T. An Empirical Evaluation of the System Usability Scale.

Die Werte des **TOAST_P (Performance)** liegen zwischen **3,80** und **5,60**, der Durchschnitt beträgt **4,89**. Die Standardabweichung von **0,54** zeigt, dass das Vertrauen in die Performance relativ einheitlich

bewertet wurde, aber im mittleren Bereich liegt. Nach Aussage der Ingenieurpsychologen der TU Darmstadt wird allgemein in der Literatur ein Wert von 5,5 als „gut“ bewertet.

Weiter bewegen sich die Werte des **TOAST_U (Understanding)** zwischen **4,75** und **6,50**, mit einem Durchschnitt von **5,73**. Die Standardabweichung beträgt **0,42**, was auf eine relativ gleichmäßige Einschätzung hindeutet. Das Verständnis des Systems ist im Vergleich zur Performance tendenziell höher.

Zur Übersicht der Ergebnisse sind in Abbildung 12: Boxplots der Ergebnisse des "System Usability Scale" (SUS) und dem "Trust of Automated Systems Test" (TOAST) aufgeschlüsselt in die Subgruppen "Performance" und "Understanding" Boxplots der drei Skalen gegeben.

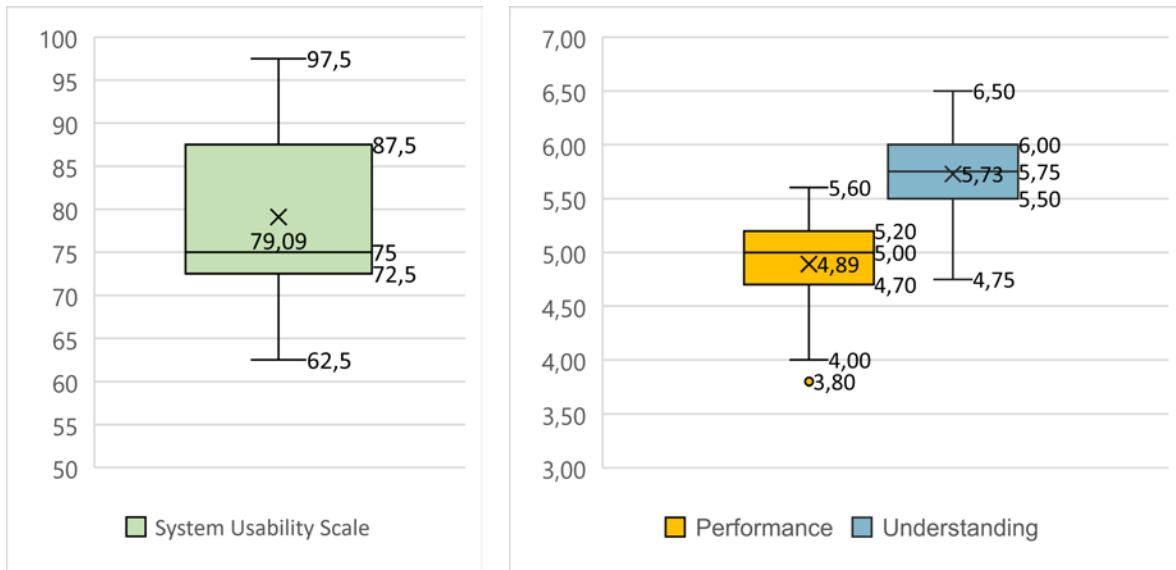


Abbildung 12: Boxplots der Ergebnisse des "System Usability Scale" (SUS) und dem "Trust of Automated Systems Test" (TOAST) aufgeschlüsselt in die Subgruppen "Performance" und "Understanding"

Zusammenfassend ist die Standardabweichung bei den SUS-Werten mit 10,57 am höchsten, was darauf hinweist, dass es größere Unterschiede in der Wahrnehmung der Benutzerfreundlichkeit gab. Im Vergleich dazu sind die Standardabweichungen bei TOAST_P (0,54) und TOAST_U (0,42) geringer, was auf eine homogenere Bewertung des Vertrauens in die Systemleistung und das Verständnis des Systems schließen lässt.

Die Studie zeigt, dass die Piloten das System insgesamt als benutzerfreundlich empfanden, und dass sie moderates bis hohes Vertrauen in die Leistungsfähigkeit des Systems bei einem hohen Verständnisgrad hatten.

Die etwas geringere Leistungsfähigkeit des Systems war dabei erwartbar, da es sich bei dem evaluierten System um einen Technologiedemonstrator mit eingeschränkter Funktionalität handelte.

Ergebnisse der Interviews

Nach Abschluss der Fragebögen wurde mit jedem Probanden ein abschließendes Interview über die Akzeptanz von KI-Algorithmen im Cockpit geführt. Innerhalb der Interviews wurden zwar vorher definierte Fragen als Diskussionsgrundlage gestellt, jedoch wurde darauf geachtet, dass die Probanden selbst Ihre Schwerpunkte setzen konnten, um ein möglichst breites Meinungsspektrum einzufangen. Nachfolgend werden die dabei gewonnenen Erkenntnisse zusammengefasst.

Die grundsätzliche Haltung zur KI im Cockpit war, dass KI Anomalien erkennen und Ratschläge geben kann, jedoch nur als beratende Instanz fungieren sollte. Die Kontrolle sollte nicht von der KI übernommen werden, sondern die endgültige Entscheidung stets beim Piloten liegen. Es wurde betont, dass durch Erfahrungen mit KI die Akzeptanz steigen kann, jedoch müssten dabei auch potenzielle Risiken berücksichtigt werden. Die Qualität des Systems hänge stark von den Entwicklern und den verwendeten Trainingsdaten ab, sodass Fehler möglicherweise unentdeckt bleiben.

Bezüglich des Mehrwerts für die Sicherheit wurde angenommen, dass KI einen positiven Beitrag leisten kann, vorausgesetzt, ihre Zuverlässigkeit ist hoch und der Einfluss der KI bleibt überschaubar. Was die Fehlertoleranz der Piloten gegenüber KI-Systemen betrifft, bestand grundsätzlich Vertrauen in diese Technologien, wobei die Erklärbarkeit der Entscheidungen als entscheidend angesehen wurde. Solange das System einen echten Mehrwert bieten würde und vom Piloten überstimmt werden könne, wurde es als nützlich und unterstützend betrachtet.

Die Bereitschaft zur Bereitstellung von Trainingsdaten war bei den Piloten hoch, sofern ihnen ausreichende Hintergrundinformationen und Erklärungen gegeben werden. Allerdings wurde darauf hingewiesen, dass falsches Lernen aus diesen Daten negative Auswirkungen haben könnte. Ein weiterer Punkt ist, dass die Aufmerksamkeit des „Pilot Monitoring“ durch eine zweite Überwachungsinstanz, wie dem Approach Briefing Assistent, gesteigert werden könnte. Die Nutzung solcher Systeme sollte jedoch auf die Einhaltung der Standardbetriebsverfahren (SOP) beschränkt bleiben.

Eine offene Frage bleibt, inwieweit der Approach Briefing Assistent die Versäumnisse der Piloten abdecken kann.

4. Veröffentlichungen, Vorträge

Im Herbst 2022 wurden die bis dato erzielten Arbeitsergebnisse auf dem „International Workshop on ATM and CNS“ (IWAC) vorgestellt. Dazu entstand eine entsprechende Publikation mit dem Titel

„Automatic speech recognition in noise polluted cockpit environments for monitoring the approach briefing in commercial aviation“

(dt. „Automatische Spracherkennung in lärmbelasteten Cockpit-Umgebungen zur Überwachung des Anflug-Briefings in der kommerziellen Luftfahrt“, die im entsprechenden Tagungsband (Proceedings) veröffentlicht wurde.

Darüber hinaus ist aktuell eine abschließende Veröffentlichung in einer begutachteten Fachzeitschrift in Arbeit. Diese soll die Aspekte Gesamtsystem, Mensch-Maschine-Schnittstelle sowie Evaluation und Zertifizierungsstrategien ergänzend zur oben genannten Veröffentlichung herausarbeiten.

Neben den schriftlichen Publikationen konnte das Konsortium auf der „Internationalen Luftfahrtausstellung“ (ILA) 2024 auf dem Stand des Bundesministeriums für Wirtschaft und Klimaschutz (BMWK) die Arbeitsergebnisse des ViCKI-Projekts einem breiten Fachpublikum vorstellen. Die Präsentation bot den Besuchern eine grundlegende Erläuterung des „Virtuellen Crewassistenten“ und gewährte Einblicke in die zwei entwickelten Anwendungsfälle: den „Approach Briefing Assistent“ und die „Flight Diversion“. Die Technische Universität Darmstadt und die Technische Universität Braunschweig trugen dabei die Hauptverantwortung für die Präsentation der Inhalte auf der Messe.

Beide Anwendungsfälle wurden im Kontext des übergeordneten Konzepts des Virtuellen Crewassistenten präsentiert. Zur Unterstützung der Einordnung wurde begleitendes Material in Form einer Folienpräsentation genutzt. Dabei kamen auch Teile der Materialien aus der abschließenden Evaluation zum Einsatz, um Einblicke in die Gewinnung der Forschungsergebnisse zu geben.

Die TU Braunschweig demonstrierte den „Approach Briefing Assistent“ in Form der entwickelten EFB-Applikation zur Vollständigkeitsprüfung von Approach Briefings mittels Spracherkennung. Besucher hatten die Möglichkeit, diesen Assistenten nach einer kurzen thematischen Einführung und mit entsprechendem Vorlagenmaterial selbst auszuprobieren und eigene Spracheingaben zu tätigen. Auf Grund der herausfordernden Akustik der Messehalle wurde neben der eigentlichen Funktionalität auch ein Demonstrationsmodus implementiert, um anhand von beispielhaften Daten den Programmablauf auch ohne Aufnahme eines eigenen Briefings zu demonstrieren.

Außerdem präsentierte die TU Darmstadt den Anwendungsfall der „Flight Diversion“. Diese EFB-Applikation soll Piloten in zukünftigen „Reduced Crew Operations“ unterstützen. Besucher konnten das System ebenfalls selbst testen, nachdem sie eine Einführung in das neue operationelle Konzept, die Funktionsweise der Cockpitassistenten und das dargestellte Szenario erhalten hatten.

Abbildung 13 zeigt den Aufbau des Anwendungsfalls „Approach Briefing“ auf dem Messestand.

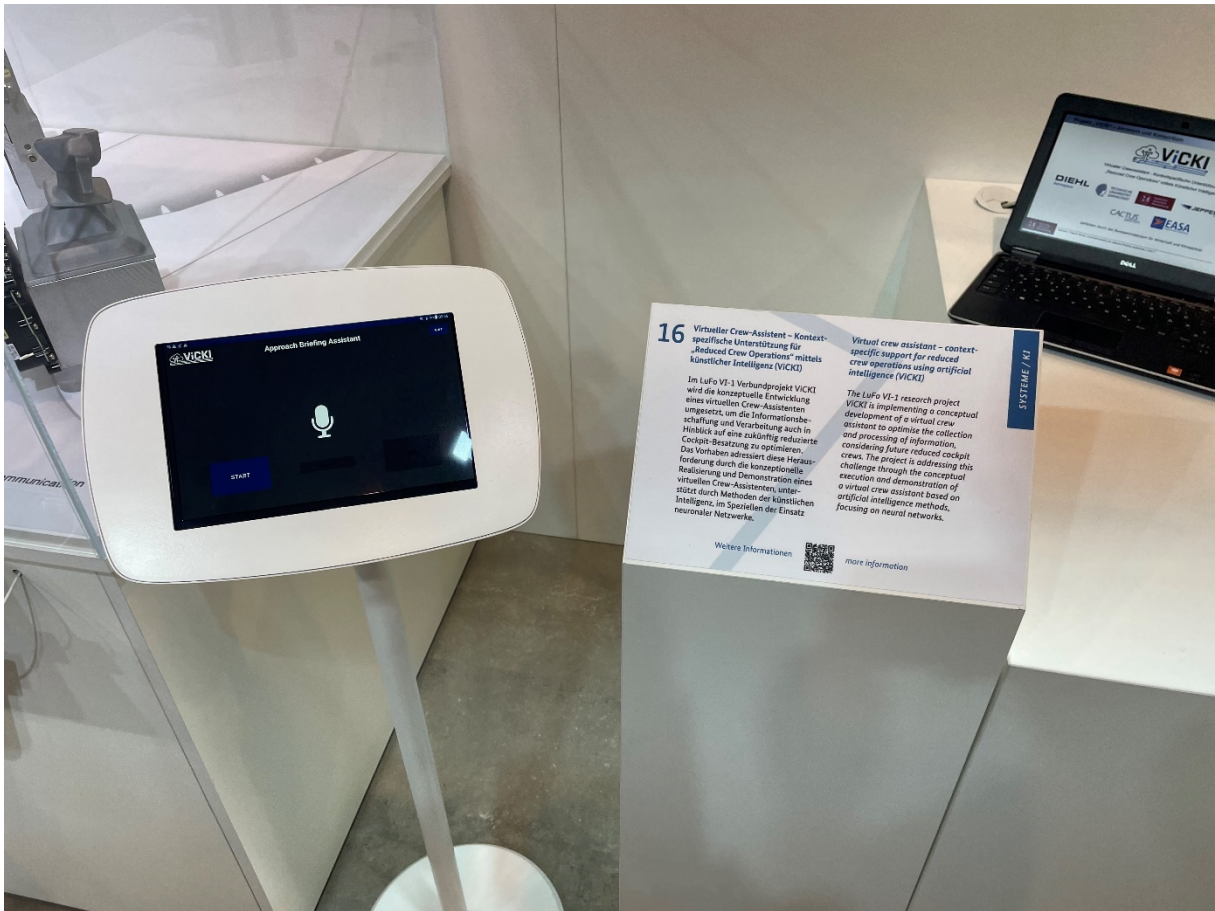


Abbildung 13: Präsentation des Anwendungsfalls "Approach Briefing" auf der ILA Berlin 2024

5. Anhang

Evaluation Study on the ViCKI Approach Briefing Assistant at the Institute of Flight Guidance, TU Braunschweig

+++ Participant Questionnaire +++

Contact information:

Institut für Flugführung
Technische Universität Braunschweig
Hermann-Blenk-Straße 27
38108 Braunschweig

Sven Bollmann	0531/391-9888	s.bollmann@tu-braunschweig.de
Jonas Füllgraf	0531/391-9873	j.fuellgraf@tu-braunschweig.de
Christian Roxlau		c.roxlau@tu-braunschweig.de

Pilot information:

Participant Number

Gender:

Licenses:

**Additional Qualifications
(Instructor, ...):**

Type ratings:

Total Flight Experience:

(hours)

Experience on A32F:

(hours)

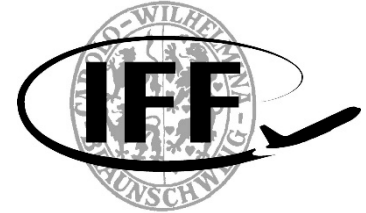
**Area of Operation
(Passenger Flight, Research, ...):**



Einverständniserklärung

zur Teilnahme an der wissenschaftlichen Studie

Evaluation Study on the ViCKI Approach Briefing Assistant



Aufklärungsbogen

Die Richtlinien der Deutschen Forschungsgemeinschaft (DFG) sehen vor, dass sich die Teilnehmer*innen an empirischen Studien mit ihrer Unterschrift explizit und nachvollziehbar einverstanden erklären, dass sie freiwillig an unserer Forschung teilnehmen. Aus diesem Grund möchten wir Sie bitten, die nachfolgenden Erläuterungen zum Inhalt der Studie zu lesen und die untenstehende Einverständniserklärung zu unterzeichnen, sofern Sie damit einverstanden sind. Ziel dieser Studie ist die Erkennung von Cockpitsprache und deren Weiterverarbeitung mit Methoden der Künstlichen Intelligenz (weiterführende Details in der Einführungs-Präsentation). Die Teilnahme an der Studie wird voraussichtlich 2,5 Stunden in Anspruch nehmen. In dieser Studie werden Sie – in der Rolle Ihrer beruflichen Qualifikation (FO/CPT) – mehrere Approach Briefings zu einer im Vorfeld geschilderten Situation durchführen. Die dabei erhobenen Sprachaufzeichnungen werden zu wissenschaftlichen Zwecken ausgewertet. Die Teilnehmenden an dieser Studie werden keinem Risiko ausgesetzt, das über die Risiken des alltäglichen Lebens hinausgeht. Sie können Ihre Teilnahme an der Studie jederzeit und ohne Angabe von Gründen abbrechen, ohne dass Ihnen dadurch Nachteile entstehen. Im Falle eines Abbruchs werden alle bis dahin erhobenen Daten gelöscht.

Erklärung zum Datenschutz

Die Datenverarbeitung dieser Studie geschieht nach datenschutzrechtlichen Bestimmungen der Datenschutzgrundverordnung (DSGVO). Die Daten werden ausschließlich für die hier beschriebenen Zwecke verwendet.

Im Rahmen dieser Studie werden folgende Daten erhoben:

- Sprachaufzeichnung des Approach Briefings im institutseigenen Cockpitsimulator
- Fragebogen zur Benutzererfahrung
- Tonaufzeichnung einer Expertendiskussion im Anschluss an die Versuche

Als personenbezogene Daten werden erhoben:

- Geschlecht
- Fliegerische Qualifikation

Vertraulichkeit

Alle im Rahmen dieser Studie erhobenen Daten sind vertraulich und werden nur in anonymisierter Form genutzt. Die personenbezogenen Daten dienen zur statistischen Einordnung der Studienergebnisse.

Aufbewahrung

Die mit dieser Studie erhobenen Daten werden an der Technischen Universität Braunschweig sowie den assoziierten Projektpartnern gespeichert.

Freiwilligkeit & Rechte der Versuchspersonen

Ihre Teilnahme an dieser Untersuchung ist freiwillig. Es steht Ihnen zu jedem Zeitpunkt dieser Studie frei, Ihre Teilnahme abubrechen und damit diese Einwilligung zurückziehen (Widerruf), ohne dass Ihnen daraus Nachteile entstehen. Wenn Sie die Teilnahme abbrechen, werden keine Daten gespeichert und alle bisher vorliegenden Daten zu Ihrer Person vernichtet. Sie haben das Recht, Auskunft über die Sie betreffenden personenbezogenen Daten zu erhalten sowie ggf. deren Berichtigung oder Löschung zu verlangen. In Streitfällen haben Sie das Recht, sich beim Niedersächsischen Datenschutzbeauftragten zu beschweren (Adresse s.u.).

Einverständnis

Ich habe die Erläuterungen zur Studie gelesen und bin damit einverstanden, an der genannten Studie teilzunehmen. Ich erkläre mich einverstanden, dass die im Rahmen der Studie erhobenen Daten zu wissenschaftlichen Zwecken ausgewertet und in pseudonymisierter Form gespeichert werden. Ich bin mir darüber bewusst, dass meine Teilnahme freiwillig erfolgt und ich den Versuch jederzeit und ohne die Angabe von Gründen abbrechen kann. Mir wurden alle Fragen bzgl. der Studie durch den anwesenden Versuchsleiter beantwortet.

Ort, Datum

Name (in Druckschrift)

Unterschrift Versuchsperson

Bei Fragen, Anregungen oder Beschwerden können Sie sich jederzeit an die Versuchsleiter wenden, die ebenfalls für die Datenverarbeitung der Studie verantwortlich sind:

Sven Bollmann: s.bollmann@tu-braunschweig.de, Tel. 0531/391-9888

Jonas Füllgraf: j.fuellgraf@tu-braunschweig.de, Tel. 0531/391-9873

Christian Roxlau: c.roxlau@tu-braunschweig.de

Bei Fragen zum Datenschutz kann auch der Datenschutzbeauftragte der TU Braunschweig kontaktiert werden:

Postfach 3329, D-38023 Braunschweig
Bienroder Weg 80, Raum-Nr. 017b,
D-38106 Braunschweig
Tel. +49 (0)531/391-7654
E-Mail: datenschutz@tu-braunschweig.de

Self-generated question Concerns attitude towards AI before interaction and experience	strongly disagree	disagree	somewhat disagree	neither disagree nor agree	somewhat agree	agree	strongly agree
I think that speech-based AI systems can add value to the Approach Briefing.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

The following questions relate to the Approach Briefing System (system) that you have just tested in the simulator. Please read the questions carefully and answer them all. Answer spontaneously and how you felt during the simulation. Please put only one cross per question, if you want to change your answer, cross out the wrong one and correct it.

Self-generated (4 questions) Concerns CS 25.1302	strongly disagree	disagree	somewhat disagree	neither disagree nor agree	somewhat agree	agree	strongly agree
The system is in a clear form appropriate to the task.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The system is in an unambiguous form.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The system enables awareness due to accomplishing the task.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The system enables to manage errors due to accomplishing the task.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

What do you base this on?

TOAST (9 questions) Concerns EFB	strongly disagree	disagree	somewhat disagree	neither disagree nor agree	somewhat agree	agree	strongly agree
I understand what the system should do.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I understand the limitations of the system.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I understand the capabilities of the system.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I understand how the system executes tasks.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The system helps me achieve my goals.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The system performs consistently.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The system performs the way it should.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am rarely surprised by how the system responds.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel comfortable relying on the information provided by the system.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

What do you base this on?

Self-generated (3 questions) Concerns EFB	strongly disagree	disagree	somewhat disagree	neither disagree nor agree	somewhat agree	agree	strongly agree
The system status is displayed in an appropriate manner to accomplish the task.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The color coding of the system messages are appropriate to accomplish the task.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The system gives appropriate feedback which entry is suspect and specifies what type of data is expected.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

What do you base this on?

SUS (10 questions) Concerns EFB	strongly disagree	disagree	neither disagree nor agree	agree	strongly agree
I think that I would like to use this system frequently.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found the system unnecessarily complex.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I thought the system was easy to use.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I think that I would need the support of a technical person to be able to use this system.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found the various functions in this system were well integrated.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I thought there was too much inconsistency in this system.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I would imagine that most people would learn to use this system very quickly.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found the system very cumbersome to use.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I felt very confident using the system.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I needed to learn a lot of things before I could get going with this system.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

What do you base this on?

Self-generated question Concerns attitude towards AI after interaction and experience	strongly disagree	disagree	somewhat disagree	neither disagree nor agree	somewhat agree	agree	strongly agree
I think that speech-based AI systems can add value to the Approach Briefing.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>