



LUND
UNIVERSITY

LUSA

SCHOOL OF AVIATION

Artificial Intelligence in Airline Pilot Training

Sezgin Kupeli

Bachelor Thesis 15 Credits

Bachelor programme in Aviation, Later Part
FLYL01

Spring Term 2025

Supervisor: Nicklas Dahlström

Artificial Intelligence in Airline Pilot Training

Sezgin Kupeli

Lund 2025

Kandidatuppsats 15 hp
Kandidatprogrammet Trafikflygare, Senare Del
FLYL01
Vårtermin 2025
Handledare: Nicklas Dahlström

Artificial Intelligence in Airline Pilot Training

Sezgin Kupeli

Number of pages: 30

Keywords:

Artificial Intelligence, Airline Pilot Training, Machine Learning.

Abstract

This thesis aims to identify current applications of artificial intelligence in airline pilot training, review existing research, evaluate potential improvements, and examine associated limitations. A literary study was conducted to achieve this. Open research was found to be limited but broad, showing a wide range of potential uses. A few airlines and training providers were found to be currently utilising AI to improve their training in various ways. They, however, only provided limited information about how it is operationalised. The study found that AI offers significant promise but faces challenges related to transparency, reliability, bias, and ethical considerations. For more widespread utilization of AI in airline pilot training, more research will likely be needed.

© Copyright: Lund University School of Aviation, Faculty of Engineering

Lund University, Lund 2025

Lunds Universitet Trafikflyghögskolan, Lunds tekniska högskola, Lunds universitet, Lund 2025.

Lunds Universitet Trafikflyghögskolan
Lunds tekniska högskola
Lunds universitet
Box 118
221 00 Lund

<http://www.tfhs.lu.se>

Telefon: 0435 – 44 54 00

Lund University School of Aviation
Faculty of Engineering
Lund University
P.O. Box 118
SE-221 00 Lund
Sweden

<http://www.lusa.lu.se>

Telephone: +46 435 44 54 00

Table of contents

1. INTRODUCTION	1
1.1 PURPOSE.....	1
1.2 QUESTION STATEMENT.....	1
2. METHOD	2
PARTICIPANT	2
3. RESULTS	3
3.1 HISTORY OF ARTIFICIAL INTELLIGENCE.....	3
3.2 OVERVIEW OF ARTIFICIAL INTELLIGENCE.....	5
3.3 EVIDENCE-BASED TRAINING.....	9
3.4 AVIATION REGULATORS AND ARTIFICIAL INTELLIGENCE	10
3.5 STUDIES ON ARTIFICIAL INTELLIGENCE IN AIRLINE PILOT TRAINING.....	11
3.6 COMPANIES THAT UTILISE ARTIFICIAL INTELLIGENCE IN AIRLINE PILOT TRAINING.....	13
3.7 INTERVIEW WITH LUFTHANSA AVIATION TRAINING	14
4. DISCUSSION.....	16
4.1 LIMITATIONS.....	20
5. CONCLUSION.....	20
6. REFERENCES.....	21

1. Introduction

Pilot training has come a long way since the Wright Brothers first took flight in 1903 (Barata, 2017). Throughout the years, as the number of flights increased, more comprehensive training programs were developed (Barata, 2017). In 1929 the first commercially-built version of a flight simulator, the link-trainer, was created to allow pilots to practice flying by instruments (Myers, 2018). As different technologies advanced so did flight simulators. Today, full flight simulators can simulate complex aircraft system and flight characteristics (Myers, 2018).

In more recent years, a new training and assessing methodology has been developed; Evidence-Based Training (EBT) (International Civil Aviation Organization [ICAO], 2013). This new framework focuses on training and assessing pilot competencies rather than the completion of a number of predictable worst-case scenarios.

Over the next 20 years, global passenger numbers are expected to increase by 3.8% per year on average (IATA, 2024b). To keep up with this growth Boeing (2024) forecasts that 674,000 new pilots will be needed between 2024 and 2043 to maintain the global commercial fleet. As commercial flights become more numerous and the need for pilots increases, improvements in the training system will become increasingly important to maintain or increase current safety standards, reduce costs and not hinder market growth.

Since the release of ChatGPT in late 2022, Artificial Intelligence (AI) has been receiving increasing attention (Marr, 2023). It has shown to be a useful tool in many areas such as engineering, business, medicine, weather forecasting, etc. (Patil et al., 2023). These areas are already seeing an increase in quality and efficiency. According to a bibliometric analysis (Roza Sh. Akhmadieva et al., 2024), research on AI in education has increased dramatically from 2004 to 2023, with a particularly dramatic increase after 2019.

Every day over 100,000 flights generate large amounts of data (International Air Transport Association [IATA], 2024a), this data can potentially be processed through AI in order to improve pilot training. The same idea could be applied to information gathered by flight simulators as well as other channels of data. AI is expected to revolutionize the world in the so-called fourth industrial revolution (Velarde, 2019). It is expected to transform virtually all sectors, enhancing productivity and enabling new business models (Velarde, 2019). As we look into the future, AI has the potential to become a powerful training tool, moving pilot training into a new era.

1.1 Purpose

The purpose of this study is to gain a broad understanding of AI, how it works, identify how it is currently being implemented in airline pilot training and explore its potential to further enhance training effectiveness, safety, and efficiency. It also seeks to identify challenges, risks and limitations with AI integration.

1.2 Question Statement

How is AI currently used in airline pilot training, what research exists on its application, and what are the key advantages and limitations?

2. Method

To gain an overview of AI research in airline pilot training the search portal LUBSearch was used to find relevant scientific articles. A search was made using the words “artificial intelligence in airline pilot training” on 20th March 2025. Non-peer reviewed articles were filtered out and no date range was specified resulting in 17 search results. After four duplicates were removed 13 scientific articles remained.

Titles, abstracts and some other parts of the articles were screened for relevance. Criteria for relevance included: Use of AI to improve pilot training or pilot selection (if the findings could be transferable to training). The latter was included due to the relevance of assessment in airline pilot training. This left a total of five articles. After the full articles were read, an additional article was excluded leaving a total of four articles.

To gain an overview of AI in general and its history, multiple search engines were used (LUBSearch, Google Scholar and Google). Due to the breadth of the topic, a concise and structured search was not possible. Multiple searches were made with various search words being used. Due to this, this part of the research has not been documented in detail and consequently cannot be accurately presented in this paper. 15 sources were used for AI history and 10 sources were used to give an overview of AI. These were selected based on relevance and credibility. Mainly scientific articles were used although some webpage-sources such as IBM (2021) and Goldman Sachs (2023) were also used due to their respective expertise and their ability to provide the most up-to-date perspectives on recent developments in artificial intelligence.

To identify companies utilizing artificial intelligence in pilot training, Google was used as the primary search engine. This approach was chosen because Google offers the most comprehensive and up-to-date coverage of company websites, news articles, and industry reports, which are often not well covered in academic databases. Multiple searches were made using phrases such as, but not limited to: “AI”, “artificial intelligence”, “airline pilot training”, “flight simulators” and “CAE”. Based on the findings, further searches were made to find more information about interesting projects.

The following companies were found to be incorporating AI in pilot training in some form: Airline Pilot Club (APC) in collaboration with Amelia, CAE, Lufthansa Aviation Training (LAT) in collaboration with 3Spin Learning, and Hinfact. Emails were sent to APC, Amelia, CAE, LAT and 3Spin Learning requesting an interview. APC, Amelia and CAE did not answer. 3Spin Learning was unable to share information due to confidentiality agreements. LAT accepted the request and an interview was scheduled. Although their project was not directly related to airline pilot training, the interview was conducted to find potential parallels. Ahead of the interview a few questions were prepared, each with follow up questions, to be used dependent on the interviewee's answers.

Participant

A one-hour semi-structured interview was conducted with Saskia Foysi, Product Manager Digital Learning at Lufthansa Aviation Training (LAT). Foysi was involved in the development of the 3spin AI Soft skills project. The interview was conducted on the 11th of April 2025 through Microsoft Teams.

Foysi was not personally working with the Evidence Based Training (EBT) methodology. She was unable to connect me with someone within LAT working on an EBT related project.

3. Results

3.1 History of Artificial Intelligence

In 1942, American Science Fiction writer Isaac Asimov published a short story called *Runaround* (Haenlein & Kaplan, 2019). This story is said to have inspired generations of scientists in the field of robotics, AI and computer science (Haenlein & Kaplan, 2019).

3.1.1 Alan Turing and the Turing Test

In 1938, British Mathematician Alan Turing published a paper in which he proposed a machine capable of computing any computable function (Turing, 1938). This would later be known as the Turing Machine (Smith et al., 2006). This paper and the Turing Machine provided the basis for the theory of computation (Smith et al., 2006).

In 1950, while he was the Deputy Director of the computing laboratory at the University of Manchester, Turing published a paper that presented a test which he then referred to as “the imitation game” (Smith et al., 2006). This would later come to be known as the Turing Test, a way of testing if machines were capable of thinking (Smith et al., 2006). This test involves a person (the interrogator) communicating with a machine and another person (Smith et al., 2006). Through conversation the interrogator needs to identify which is human and which is machine (Smith et al., 2006). The idea behind the test is that if a machine is indifferentiable from a human, it should be considered intelligent (Smith et al., 2006).

Turing's hope was that the test would be passed by the year 2000 (Smith et al., 2006). In a very recent study (Jones & Bergen, 2025), which is yet to be published at the time of writing this paper, ChatGPT 4.5 managed to pass the Turing test 73% of the time while LLaMa-3.1-405B managed to pass 56%, ELIZA 23% and GPT-4o 21% of the time. Turing is often referred to as the father of AI and the father of computer science (Cooper & J Van Leeuwen, 2016).

3.1.2 Dartmouth Summer Research Project on Artificial Intelligence

Later, in 1956, Marvin Minsky and John McCarthy hosted the Dartmouth Summer Research Project on Artificial Intelligence (DSRPAI), an eight-week workshop funded by the Rockefeller Foundation (Haenlein & Kaplan, 2019). It marked what would later be known as the AI Spring, a period of increased funding and optimism. Its purpose: to gather researchers from different fields and create a new research area aimed at building machines capable of simulating human intelligence. It was here the term Artificial Intelligence was officially coined (Haenlein & Kaplan, 2019). The two following decades saw significant progress in AI research (Haenlein & Kaplan, 2019).

3.1.3 Artificial Intelligence Winter

However, in 1973, the U.S. Congress became very critical of the high spending on AI research (Haenlein & Kaplan, 2019). The same year, British mathematician James Lighthill published a report questioning the optimism surrounding AI research (Lighthill, 1973). In response to this, the British government ended support for almost all universities in the country with the U.S. government following suit (Haenlein & Kaplan, 2019). This marked the beginning of the AI winter, a period where a lot of research and many projects were shut down due to reduced funding.

According to Haenlein & Kaplan (2019) one reason for the lack of progress in the field of AI was because of the way early AI systems such as the programs ELIZA and the General Problem Solver tried to replicate human intelligence. They were so called Expert Systems that had a top-down approach, using a series of if-then statements. Expert Systems performed impressively well in areas with specific set rules. For example, IBM created a program called Deep Blue, that in 1997 managed to beat World Champion chess player Gary Kasparov (IBM, n.d.). However, when faced with other problems expert systems are unable to adapt or learn. These situations require a bottom-up approach, to learn the rules from the data itself. Haenlein & Kaplan (2019) claimed that (based on their definition of AI) because Expert Systems lack the ability to learn from data and adapt, they are not true AI.

As early as the 1940s Canadian psychologist Donald Hebb developed a theory of learning known as Hebbian Learning, this method replicates the way neurons function in the human brain (Huyck, 2021). This would later develop into Artificial Neural Networks, a bottom-up approach. However, research on this stagnated in 1969 when Marvin Minsky and Seymour Papert showed that the available processing power at the time was insufficient to handle that kind of network (Minsky & Papert, 1969).

Since then, computing power has increased significantly. Moore (1965) observed that the number of components (such as transistors) that could be placed on a single integrated circuit was doubling roughly every year. He predicted that this would continue for the foreseeable future. Moore also anticipated that this trend would drive major advances in computing, miniaturization, and new applications for electronics. Later, this law was modified to doubling every two years (Zhu et al., 2023). His insight, now known as Moore's Law, became a guiding principle for the semiconductor industry and a catalyst for the rapid evolution of modern technology (Flamm, 2018).

As the years passed by, computing power increased, and Artificial Neural Networks returned in the form of Deep Learning (Haenlein & Kaplan, 2019). In 2015 Google used this method to create a program called AlphaGo that was able to beat a world champion in a game called Go, a significantly more complex game than chess (Google DeepMind, n.d.). Most programs we know today as AI are driven by Artificial Neural Networks and Deep Learning (Haenlein & Kaplan, 2019).

Throughout its history, AI has seen both ups and downs as optimism and hope have attracted large amounts of investments followed by significant drops in funding (Francesconi, 2022). Figure 1 below shows how interest in the field has varied throughout the years.

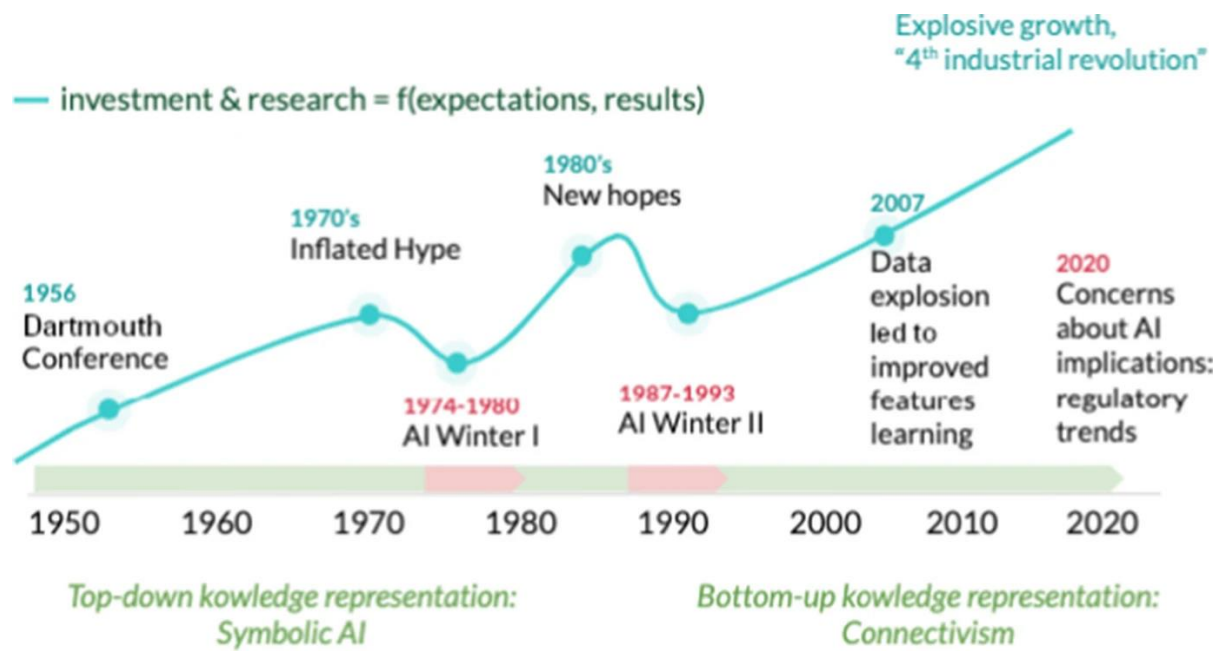


Figure 1. The ups and downs of AI investment and research, (Francesconi, 2022)

Throughout its history, experts have predicted that artificial general intelligence (AGI), AI that exhibits human-like intelligence and adaptability across diverse tasks (Raman et al., 2025), was just a few years away (Haenlein & Kaplan, 2019). However, any time AI mastered something the definition of AI or of intelligence was changed to exclude the latest AI capability. This phenomenon is called the “AI Effect” (Haenlein & Kaplan, 2019).

3.2 Overview of Artificial Intelligence

According to Kaplan & Haenlein (2019) AI is “a system’s ability to correctly interpret external data, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation”. In a more fundamental but also vague sense, AI is “a broad field of computer science concerned with building smart machines capable of performing tasks that typically require human intelligence” (Sarker, 2022). In this chapter we will explore some fundamental concepts related to AI to gain a broad, surface level understanding of how it works.

3.2.1 Expert Systems

According to Durkin (1990) Expert Systems are programs designed to model the problem-solving ability of human experts by modelling their knowledge, reasoning, conclusions and explanations. They function based on explicitly programmed rules and knowledge bases. They require rich datasets containing only relevant knowledge and well-focused tasks. One of the most common ways of representing knowledge in Expert Systems is through rules, which are represented by conditions; if-then statements. This presents knowledge as cause and effect, similar to how human experts would formulate their knowledge. These systems then reason primarily through two inference techniques: backward chaining or forward chaining. Backward chaining relies on having a goal or hypothesis and then testing

new data against preexisting knowledge (datasets). Forward chaining involves collecting new data and attempting to infer information from it based on preexisting knowledge (datasets). Both inference methods can be integrated, for example forward chaining can be used to produce a hypothesis and backward chaining can be used to verify it. A major advantage of Expert Systems is that they can explain how they reached their conclusions, providing transparency. They can easily be improved and adapted by adding, deleting or updating rules based on new knowledge or recent trends (Sarker, 2022).

3.2.2 Machine Learning

Machine Learning, also known as data-driven learning, is a subset of AI. By learning from data or experience, it automatically builds analytical models (Sarker, 2022). Various types of machine learning algorithms exist, such as supervised, unsupervised, semi-supervised, and reinforcement learning (Sarker, 2021). The primary difference between these algorithms is the way in which they learn from a dataset to create their models (Trezza et al., 2024). Machine Learning models usually consist of a set of rules, procedures or transfer functions that are used to identify patterns of interest or anticipate future behaviour. By analysing historic data, models are built which can then be used with new data to make predictions about future outcomes, which can for example help businesses assess risk.

3.2.3 Artificial Neural Networks

Artificial Neural Networks (ANN) are Machine Learning programs or models that mimic the way biological neurons work in a brain (IBM, 2021). One example of ANNs is feedforward neural networks, also known as multi-layer perceptrons (MLPs). These consists of layers of Artificial Neurons (also called nodes); an input layer, an output layer and one or more hidden layers in between. Each node is connected to other nodes. The nodes have certain weights and thresholds (also known as bias), these depend on the importance of given variables (inputs). A nodes weight determines how much influence its output has on the next node. Its threshold is how much input is required for it to activate. There are numerous types of ANNs such as Convolutional neural networks (CNNs) and Recurrent neural networks (RNNs) but due to the scope of this paper they will not be covered. ANNs require datasets and learning algorithms to create models.

3.2.4 Deep Learning and Black Box Artificial Intelligence

Deep Learning is a subset of Machine Learning (Sarker, 2022). When an ANN has more than one hidden layer (more than three total) it is considered a deep learning algorithm. It utilizes these multiple layers and processing to build data-driven models. An illustration of an ANN with deep learning can be seen below in figure 2.

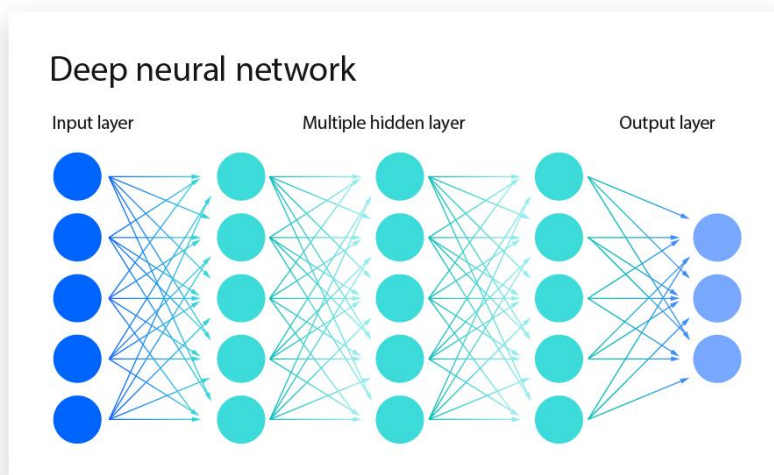


Figure 2. Deep neural network (IBM, 2021)

A problem that arises with deep learning is a lack of transparency (Kosinski, 2024). Although the input and output are visible, due to the complexity of the “hidden layers” the way in which the system achieves the output is unknown, this kind of AI system is referred to as a black box AI. Many of the most advanced machine learning models available today, such as large language models like OpenAI’s ChatGPT and Meta’s Llama, are black box AIs. Due to the large datasets they were trained on and the complex deep learning processes they use, even their own creators do not fully understand how they work. The opposite of black box AI is white box AI, also known as explainable AI (XAI) or glass box AI. Expert systems are usually considered XAI. According to Kosinski (2024) researchers are working on increasing the transparency of black box AI systems.

3.2.5 Generative Artificial Intelligence and Hallucinations

Generative AI is a computational technique that is capable of generating seemingly new content such as text, images or audio from training data (Feuerriegel et al., 2023). It draws upon patterns and relationships found in training data. Deep neural networks are particularly well suited for these kinds of models. A popular example of generative AI is OpenAI’s ChatGPT (Feuerriegel et al., 2023). GPT (generative pre-trained transformer) being a popular family of LLMs (large language models).

Industry reports predict that generative AI could raise global gross domestic product (GDP) by 7% while the equivalent of 300 million full-time jobs could be lost to automation (Goldman Sachs, 2023).

There is however a major limitation with generative AI. Due to their machine learning models relying on probabilistic algorithms for making inferences, they generate the most probable response, not necessarily the correct one (Feuerriegel et al., 2023). This means that they can produce misinformation and deceive users. This kind of behaviour is called hallucination.

3.2.6 Deep Learning Methods

Supervised learning is a task-driven strategy, a particular goal is desired. A dataset with labelled data is provided to train the algorithm to classify data or forecast outcomes (Sarker, 2022). The most common uses of supervised learning models are classification and regression. Classification involves predicting a label, such as identifying spam email, while regression involves predicting a quantity, such as the price of a house.

Unsupervised learning is a data-driven method where patterns, structures, or knowledge are derived from unlabelled data (Sarker, 2022). The most common tasks associated with Unsupervised Learning include clustering, visualization, dimensionality reduction, finding association rules, and anomaly detection. A clustering model could for example be used for identifying distinct patient subgroups (Trezza et al., 2024). Unsupervised Learning is a useful tool for complex data processing tasks where manual labelling is not possible due to the complexity of the dataset.

Semi-supervised learning is a hybrid approach; a mix of Supervised and Unsupervised Learning (Sarker, 2022). This method uses both labelled and unlabelled data to train models. This kind of learning can be useful for situations where data must be labelled automatically without human interaction.

Reinforced learning like Unsupervised Learning is trained on an unlabelled dataset (Sarker, 2022). Reinforced learning, however, utilizes a feedback loop that allows the algorithm to assess its own output. By assessing whether its solutions effectively solve a specified problem it learns through trial and error. This type of system is capable of perceiving and interpreting its surroundings as well as taking actions based on this. Therefore, it can be considered an environment-driven approach.

Based on the required task, different types of data-driven models can be built (Sarker, 2022). Common examples include analytical methods such as descriptive (what happened in the past?), diagnostic (why did it happen?), predictive (what will happen in the future?), and prescriptive (what action should be taken).

3.2.7 Fuzzy Logic

In contrast to the rule-based expert systems described earlier in this chapter, another possible approach is a fuzzy logic-based approach (Zadeh, 2008). While rule-based systems work in binary (1s and 0s, yes or no) fuzzy logic creates precise logic of unclear, incomplete or conflicting information, similar to how humans' reason. It creates scales to allow reasoning between the values of binary logic. For example, instead "is it hot or cold?" where a threshold temperature would need to be set, fuzzy logic could consider the specific temperature. It allows values between 0 and 1; it allows things to be partly true or partly false.

3.2.8 Artificial Intelligence Bias and Ethics

A challenge with AI algorithms is potential AI bias caused by unrepresentative data, human prejudices in data labelling, and algorithmic design choices (Wajid et al., 2024). These biases can manifest as selection, measurement, or algorithmic bias, often leading to unfair or discriminatory outcomes in areas like healthcare, hiring, and criminal justice. According to Wajid et al. (2024) these biases not only perpetuate social inequalities but also erode public trust in AI systems.

There are several strategies for addressing AI bias, including enhancing data quality and diversity, implementing fairness-aware algorithms, and conducting regular audits with bias detection tools (Wajid et al., 2024). Further mitigation of AI bias can be achieved by assembling diverse, multidisciplinary teams, using transparent and explainable AI systems where possible, and developing robust ethical guidelines and regulatory frameworks.

While addressing bias is crucial to ensuring fairness in AI systems, it represents just one part of the broader topic of AI ethics (Haenlein & Kaplan, 2019). The Trolley Problem is a popular thought experiment where a choice needs to be made between not acting and letting many people die or acting and killing other, albeit fewer, people. Scenarios like this need to be accounted for and when things do go wrong, the question of liability becomes highly relevant. How do you integrate ethical considerations into the advancement of AI research and technology to ensure that innovation progresses responsibly and delivers meaningful benefits to society as a whole? These are just some of the questions that researchers, developers, end users and regulators need to consider.

3.3 Evidence-based Training

In 2007, at an IATA meeting on training and qualification, stakeholders expressed the need for a strategic review of airline pilot training (ICAO, 2013). They felt that developments in aircraft design and reliability, as well as the rapidly evolving operational environment were not being adequately considered from a human factors perspective.

In response to this, IATA facilitated the creation of an international working group consisting of Civil Aviation Authorities, academic institutions, aircraft original equipment manufacturers, airlines, international organizations, pilot representative bodies and training organizations (ICAO, 2013). This group conducted a strategic review of airline pilot training, starting with recurrent training. Based on this review, the group created a new methodology for airline pilot training: Evidence-based Training.

The philosophy of EBT is to identify, develop and evaluate the competencies required to operate safely based on relevant threats identified in accidents, incidents, flight operations and training (ICAO, 2013). The idea behind the core competencies is that they are transferable. If trained in one scenario, they can be applied to others. There are eight core competencies:

- application of procedures
- communication
- aircraft flight path management, automation
- aircraft flight path management, manual control
- leadership and teamwork
- problem solving and decision making
- situation awareness
- workload management (ICAO, 2013)

Each core competency has a set of behavioural indicators which are used to assess their associated competency (ICAO, 2013). For example, the behavioural indicators for application of procedures are:

- identifies the source of operating instructions
- follows SOPs unless a higher degree of safety dictates an appropriate deviation

- identifies and follows all operating instructions in a timely manner
- correctly operates aircraft systems and associated equipment
- complies with applicable regulations.
- applies relevant procedural knowledge (ICAO, 2013)

Before the introduction of EBT, international standards for airline pilot training evolved from the evidence of accidents involving early generation jets (IATA, 2014). The belief was that repeated exposure to “worst case” scenarios would provide effective training for pilots. Over time however, novel events resulting in serious occurrences were simply added to an increasingly crowded list of manoeuvres and training scenarios. This eventually led to a “box ticking” approach to training rather than focusing on enhancing the overall abilities of the pilots undergoing training. EBT shifts the focus of the training towards competencies and training programs are adapted based on derived threats from recent evidence from multiple sources.

3.4 Aviation Regulators and Artificial Intelligence

The European Union Aviation Safety Agency (EASA) and the Federal Aviation Administration (FAA) are considered the most influential aviation regulators in the world (Tao et al., 2014). Their regulations often set benchmarks for and influence aviation authorities worldwide. Consequently, their approaches to artificial intelligence are likely to shape the industry as a whole. Both regulators have published documents outlining their strategies for integrating artificial intelligence into aviation and informing stakeholders of their intentions.

3.4.1 EASA Artificial Intelligence Roadmap 2.0

The Artificial Intelligence Roadmap 2.0 defines EASA’s strategy for the safe and trustworthy integration of artificial intelligence in aviation (EASA, 2023). The roadmap identifies several primary challenges, including ensuring public confidence, establishing robust certification for advanced automation, integrating ethical considerations such as transparency and fairness, and developing additional processes and standards to maximize AI’s safety benefits.

EASA’s stance appears to be human-centric, balancing a desire to accelerate investments and promote development but at the same time ensuring safety and ethics (EASA, 2023). In alignment with this, a regulatory proposal was published by the European Commission in April 2021, the EU AI Act. This Act is a legal framework that entered into force on 1st of August 2024 and will be fully applicable from 2nd of August 2026 (European Commission, 2024).

The EU AI Act includes a number of key provisions (EASA, 2023). It introduces a risk-based approach, defining four levels of risk for AI systems: minimal, limited, high and unacceptable risk, each with different levels of obligations (with the exception of unacceptable risk for obvious reasons). It mandates transparency (e.g., users must be informed when interacting with AI systems like ChatGPT), requires high-quality and unbiased training data, and enforces strict data protection. The Act also stipulates that appropriate human oversight must be maintained, particularly for high-risk AI applications.

3.4.2 FAA Roadmap for Artificial Intelligence Safety Assurance

The FAA has developed this roadmap to establish the guiding principles for assuring the safety of artificial intelligence in aviation and to establish priorities and plans for its safe introduction (FAA, 2024). In contrast to EASA's roadmap, the FAA's roadmap does not address ethical aspects of AI, stating that it is outside the scope of the document.

The FAA's strategy for integrating artificial intelligence in aviation follows an incremental, risk-based approach (FAA, 2024). The agency prioritizes the introduction of AI in lower-risk applications first, allowing for the accumulation of operational experience and the adaptation of safety assurance methods. Over time, lessons learned from these initial implementations are intended to inform the safe expansion of AI into higher-risk, safety-critical applications. The initial focus is on leveraging AI to enhance safety, rather than other uses such as those providing efficiency gains. They also highlight the risk in personifying AI. They claim that it can erode safety as it creates ambiguity on the assignment of responsibility for safe operations.

The FAA also acknowledges the need for in house expertise (FAA, 2024). Due to the speed at which AI research is expanding the FAA cannot rely on conventional aviation experience. Among other roles the FAA has hired a Chief Scientific and Technical Advisor for AI/ML. This will allow them to oversee the safe introduction of AI and foster continuous learning.

3.5 Studies on Artificial Intelligence in Airline Pilot Training

3.5.1 Application of Big Data and Artificial Intelligence in Pilot Training a Systematic Literature Review

This paper is a systematic literature review of 10 papers related to Big Data and AI, primarily in airline pilot training (Shaker & Al-Alawi, 2023). The reviewed papers were from between 2017 and 2022, seven of them were related to commercial aviation and three of them to military aviation.

The authors of this paper found that there is a lack of Micro-Adaptive learning in airline training programs (Shaker & Al-Alawi, 2023). Micro-Adaptive learning adapts training programs to individual pilots learning styles and abilities on a micro level in real-time. This allows instructors to focus on pilots' weaker areas rather than areas they have already demonstrated competency and proficiency in. They argued that pilots are highly motivated so by allowing them to see their own data they would be driven by self-discovery and a desire to improve.

When pilots undergo training a large amount of data is generated containing dormant patterns capable of revealing operational deficiencies (Shaker & Al-Alawi, 2023). They recommend future research focuses on utilising AI to process this data to generate real-time feedback, customised recurrent training, cognitive based assessments during training and micro-adaptive training. They found that most studies use single methods in exploring Big Data and AI applications, they argue that multi-methodological approaches would be more appropriate for future research.

3.5.2 The intelligent forecasting approach of learning curve in fits training of student pilot

In this study, machine learning was utilised to forecast student pilots learning-curves (Huang et al., 2019). This allows training programs to be adapted to the specific learning-curve of individual student pilots. Linear regression and data collected from flight academics were used to build an intelligent system. This was tested on 24 students in a recreation pilot program at an aeronautical academy in Taiwan. The program contains 12 courses and 30 flight hours as standard.

The study showed that the model could be used for prediction but requires fine-tuning as there are some deviations (Huang et al., 2019). According to the authors, these deviations were likely due to the small sample size and the difference in learning-curves between the different courses.

3.5.3 AI semantic knowledge base for Competency-based training

In this study AI was used in the form of a semantic reasoner to help develop an EBT Ontology (EBTOnto) (Dapica & Peinado, 2021). A semantic reasoner is a software tool designed to infer implicit knowledge by applying defined rules to explicit knowledge (Carbonaro, 2022). The authors intention with EBTOnto was to make it easier to reuse existing pilot training scenarios (used in simulator training) by mapping out what EBT core competencies they train. Furthermore, it incorporated a large amount of aviation incident data (37,568 incident reports) from the U.S. National Aeronautics and Space Administration (NASA) Aviation Safety Reporting System (ASRS). These reports were classified based on what lack in core competencies were identified to be likely contributing factors. This facilitated the development of new training scenarios. It also helped identify potential combinations of scenario elements that are related to errors in non-technical skills in real environments. Ultimately this could help verify training programs by quantifying the prevalence of real-world competency deficiencies allowing a direct comparison with what is being trained.

The focus of this study was on the following non-technical competencies: communication (COM), problem solving and decision making (PSD), situational awareness (SAW), and workload management (WLM). However, leadership and teamwork (LTW) could not be mapped because it was not made explicit in any of the concepts represented in the taxonomy used in the ASRS database.

The main role of AI in this study was to help determine the consistency of the ontology, it also classified the incidents, establishing whether they corresponded to valid training scenarios in accordance with the non-technical core competencies considered in the study.

In their conclusion they stated: “The use of EBTOnto in future intelligent systems opens an unexplored path for the customization of competency-based flight training scenarios, with results that can be understood and explained by the flight instructors who must apply them.” (Dapica & Peinado, 2021, p. 10).

3.5.4 Pilot Selection in the Era of Virtual Reality Algorithms for Accurate and Interpretable Machine Learning Models

Although not directly training related, this study (Ke et al., 2023) was considered relevant due to the importance of assessment in pilot training. Also, according to Ke et al. (2023) the study’s virtual reality simulation platforms and algorithms could also be used for pilot training. They argue that machine learning algorithms exhibit high predictive powers

because they can reflect complex interactions and non-linearities thanks to their ability to handle a large number of predictor variables. In the future these algorithms could include inputs such as personality, cognitive task performance, electroencephalography (EEG), heart rate, eye movements, and flight dynamics.

In order to identify and measure differences between experts and novices, 46 participants flew a traffic pattern in a flight simulator using a virtual reality headset. 23 of these participants were considered experts and 23 novices. Eye movement and flight data was collected from the flight, 8 metrics from eye tracking and 21 from flight data. These metrics were then fed into a machine learning model to infer relevant differences between expert and novice pilots.

They found that a support vector machine (SVM) algorithm performed better than other algorithms commonly used in similar studies in the past. SVM algorithms are supervised machine learning methods that help classify data (IBM, 2023). In this study the SVM algorithm proved to be robust, highly accurate and was able to mine a large amount of latent knowledge from data, meaning it could even be effectively used on smaller datasets. However, SVMs can lose performance due to data noise (irrelevant data in the dataset). To address this, the SVM algorithm was complemented by a feature that derived a mutual information coefficient (MIC). MICs help rank the importance of different features by quantifying how dependent two variables are to each other (Kinney & Atwal, 2014).

Ke et al. (2023) recommend that future research should investigate the use of other inputs such as heart rate. They argue that this increase in multimodality could lead to a more accurate assessment taking into account more of the complexities of human performance. They added that future research could also consider more complex neural networks, although interpretability could be an issue.

3.6 Companies That Utilise Artificial Intelligence in Airline Pilot Training

3.6.1 CAE

CAE is a Canadian company that manufactures simulators and provides training, including pilot training (CAE, n.d.-b). They use AI to allow their systems to automatically learn and improve from continued use and experience (CAE, n.d.-a). One example is their “CAE Trax Academy” which functions as an overarching training framework that divides the training process of pilots into three parts: learn, practise and perform (CAE, n.d.-c).

The learn stage includes courseware and virtual reality (VR) technology to help the student learn a procedure. It also utilises a feature called Adaptive Learning to capture student performance data and monitor competency acquisition (CAE, n.d.-c). From this, automated AI-based personal recommendations are provided to enhance the learning process. The student then practises to consolidate what was learned through a VR system called CAE Sprint. A virtual coach provides immediate, actionable intervention to support progress. Performance is measured from the sessions, allowing trends to be monitored, and gaps identified. Finally, the student performs in a high-fidelity simulator. Showing what they have learned and consolidated in the earlier stages. This learning management system is now being used by the U.S. Air Force.

3.6.2 Airline Pilot Club

The Airline Pilot Club™ (APC) is an Irish company that serves as an international platform supporting aspiring and experienced pilots through partnerships with airlines, flight schools, and training organizations (APC, n.d.). APC utilises SoundHoundAI's AI agent platform Amelia (APC, 2025; Amelia, n.d.). According to APC (2025) this AI agent has already been used with partner companies to make recruitment more efficient, enhance training by enabling them to realise the efficiencies provided by EBT and increase operational safety and efficiency. In the future APC (2025) aims to use this AI agent to introduce:

Enhanced predictive analytics for pilot performance and training needs. Expanded scenario-based training scenarios using AI-driven adaptive learning. New applications for non-pilot aviation roles, including cabin crew and maintenance staff. Deeper integration with industry standards like CBTA and EBT frameworks, ensuring regulatory compliance and operational excellence. (APC, 2025)

3.6.3 Hinfact

Hinfact is a French software company that utilises human factors research and technological innovation to develop training tools (Hinfact, n.d.-a). Among other things, they provide automated and adaptive training management systems through the use of AI (Hinfact, n.d.-b). These systems include automatic detection of observable behaviours based on simulator data, including eye-tracking data, enabling automatic grading based on predefined rules (Hinfact, n.d.-c). This allows for objective insights, providing greater consistency. This data can further be used to provide analytics, allowing tracking of long-term progress.

AI is also used to automate workflows (Hinfact, n.d.-c). For example, training programs can be adapted to individual trainees, and training schedules can be automated, with monitoring of resources such as simulators, aircraft, and instructors.

3.6.4 Lufthansa Aviation Training

Lufthansa Aviation Training (LAT) is a German training provider for over 250 airlines including the Lufthansa Group (LAT, 2024). They have been integrating AI into their training programs in the context of various pilot projects (LAT, n.d.). One example is by extending a LLM from GPT by adding a context layer, consisting of data relevant to training and flight operations such as operative manuals. This can then be accessed in simulator training scenarios as AI support via a chat or voice interface.

Moreover, they have been working on a soft skill, de-escalation training prototype in collaboration with 3Spin Learning, a company specialized in utilising VR and augmented reality (AR) and AI to digitize learning (LAT, 2024). This is covered below in 3.7.

3.7 Interview with Lufthansa Aviation Training

Foysi is involved in the development of most digital trainings within LAT, from web-based training to VR and Extended Reality (XR).

In this field, there is a lot of focus on the utilization of AI. One use is content creation, as it makes the creation of pictures, videos, etc. easier. Traditionally, soft skill training has been conducted either by watching a video in a classroom and then using this video as a basis for de-briefing or through role-play. Due to the lack of immersion with typical role-playing

exercises trainees may not behave very similarly to how they would in a real scenario. This is where AI/XR has potential to facilitate better training. Lately LAT has been working on implementing AI more directly with training. They are currently developing a prototype with 3Spin Learning which utilizes a GPT interface inside a unity engine (a game engine). The main advantage they see with this is a more personalized learning experience for the trainee. The trainee wears VR goggles and interacts with an AI-supported avatar to train soft skills. This creates a realistic yet safe environment in which trainees can try different communication strategies and see their impacts.

This allows trainees to try different strategies which are recommended in dealing with difficult situations, such as having to deal with passengers that refuse to follow rules. Foysi highlights the importance of practice, as many of these situations don't always have a "right or wrong" way of dealing with them. Although this project was initially small, a practical approach using trial and error, it shows potential for future development with the AI not only reacting to the trainees' spoken words but to other data such as tone of voice, and position/distance. This data could be used to adapt the avatars behaviour to the trainee's reactions and behaviour.

This data could also be generally useful for suggesting training that is specifically aligned with trainees' competency goals. If and which data may be collected and used always depends on the individual requirements of the respective airlines.

The current version only uses language spoken by the trainee within the current session. It does not utilize data from previous sessions or other sources of data.

Initially this project was aimed at being used as a basis for de-briefing in a classroom, but now we have the potential to provide feedback through an AI-driven instructor. This means that the practice could be done completely without the use of a human instructor. It is important that the AI is given accurate and reliable models or tools in order to give consistent and valid feedback. The AI-driven instructor is currently in the "testing and checking phase".

The AI analyses what the trainee says and compares it to a methodology that's given in a briefing and checks whether what was said matches certain criteria. The more context you give the AI the better. For example, if six-step approach is recommended for a certain scenario, it can check if you apply those steps. They found this approach to be the best due to its transparency.

This could lead to cost saving as less time needed in classrooms could reduce the need for location infrastructure as well as trainers. The most expensive aspect of development is creating the virtual environment, the avatars and their animations (such as gestures). Another significant cost is the hardware, the VR goggles. Although the initial cost is high, once purchased they can be used for a long time. After these initial costs, it is relatively easy to add more avatars with different briefings and create new scenarios. Meaning it is adaptable and scalable.

As this is quite new technology and many people do not have their own VR goggles at home, there could be a need for some level of location infrastructure and technical support in its initial roll out. This could create a bottleneck and reduce cost saving early on.

Another project that LAT is currently working on is the D-CEET (Digital Cabin Emergency Evacuation Trainer) project in collaboration with ETH Zürich which is more of a research-based project utilizing eye-tracking, AI and VR. Foysi was not personally involved in this project and was unable to arrange an interview with anyone involved in this project.

LAT currently do not have any data that shows any increase in learning or retention rate compared to traditional methods as they are still in the prototype and test phase. The prototype had only been tested on trainers and training managers at the time of the interview. Foysi referred to a study by PWC (2021) demonstrating that soft skill VR training yields better results compared to web-based or classroom training.

One thing LAT learned from the internal tests is that the VR soft skill training is surprisingly efficient even when the surroundings are not perfectly realistic due to the focus being on the interaction. Response time, however, is important. A lag in the avatar's response takes away from the immersion. So, the focus will be on improving the response time and ensuring a good internet connection in order to ensure a smooth experience. Overall, it appears to be motivating as trainees get the opportunity to try different methods without putting themselves in danger.

This prototype is currently not planned for airlines outside of the Lufthansa Group. It has been created by LAT, and it might serve as the basis for adaptations for different airlines within the group.

The biggest challenge in the development of the soft skills training prototype was to make the AI stay in its role and not be too independent. To create a realistic scenario but still have some sort of control so it goes in the right training direction. The available data set was small for this prototype meaning the briefing given to the AI needs to be particularly precise so the training does what it should.

4. Discussion

Airline pilot training today is widely being conducted through the EBT methodology. This utilises recent data from multiple sources to adapt training programs to “the threats of today” while focusing on the required core competencies to mitigate those threats. AI could potentially process this data very effectively as well as improve training in other areas. Analytical AI for example could be used to “diagnose” the root cause of performance deficiencies; it could also forecast potential future problems. There are however a few flaws that must be considered: black box AI, AI hallucinations, AI bias and ethical considerations. Aviation regulators must balance the potential benefits of AI with its inherent risks, creating regulations that help AI integration and development while protecting those subject to their regulations.

History

AI research has seen periods of optimism and progress followed by periods of pessimism and a reduction in research (Francesconi, 2022). This was usually due to the AI effect and overly ambitious goals. However, we have seen a consistent increase in computing power in accordance with Moore's law (Moore, 1965). This increase in computing power allows for more complicated and demanding algorithms to be developed, as well as larger datasets to be processed. Whether we have reached the limit like in 1969 when research on ANNs stagnated due to limited computing power (Minsky & Papert, 1969) is not clear based on the papers included in this study. If history is to repeat itself, we might see another downturn. Whether this will happen and if so when is however not clear.

Generative AI and Hallucinations

AI hallucinating is similar to humans drawing conclusions based on correlation rather than causation. AI simply uses probabilistic algorithms to make predictions based on historic data regardless of whether there is evidence of causation (Feuerriegel et al., 2023). This could reduce the efficiency of training if training scenarios are made based on this. Similarly, AI may erroneously classify certain behaviours as negative due to hallucinations, potentially leading to misguided conclusions and interventions. This could affect the individuals being trained or examined unfairly. Due to these limitations, generative AI should be used with caution. If used, it should include strong safeguards capable of limiting its agency and appropriate monitoring. A threat matrix could be used to assess the potential risk in implementing it in different areas. The way in which AI is being used by LAT seems like a good example of a low-risk application.

Future research could explore algorithms that are capable of reasoning in a way that infers causality, judges where it lacks sufficient evidence and limits itself appropriately. Until an acceptable level of accuracy and reliability has been achieved, simpler versions of AI such as expert systems could be used where possible.

Black box AI

Incorporating complex deep learning algorithms in airline pilot training presents significant challenges due to their lack of transparency (Kosinski, 2024). Airlines function within highly complex and dynamic environments, where changes in one area can unpredictably influence multiple others, sometimes with severe or even fatal consequences. If the performance of flight crew is altered in any way by changes in training, safety margins could be reduced. The inability to comprehend how AI systems arrive at their conclusions renders it impossible to verify or validate their decision-making processes.

In the event of errors or adverse outcomes, there are also issues with liability. Determining responsibility is complicated when the underlying processes are opaque, raising critical questions about who should be held accountable and who would be willing to assume responsibility for decisions made by systems that cannot be fully understood.

AI Bias and ethics

Aviation training is an important part of what makes aviation so safe. Training scenarios have long been based on data, initially on historic worst-case scenarios but lately with the introduction of EBT, on more detailed and up to date information (IATA, 2014). Corrupted data, whether intentional or not, can therefore potentially undermine the basis of the training itself (Wajid et al., 2024). There could be bias in the data or inherently in the way an algorithm works. It could pick up certain trends more easily than others, the same way certain threats may appear more salient than others to humans despite not having a higher threat level. This bias could affect the efficiency of training or even directly reduce safety margins. Bias could also affect individuals, as mentioned earlier under the hallucination part of the discussion, as behaviours exhibited by certain individuals could potentially be interpreted in an unfair way. There are numerous ways to minimize bias (Wajid et al., 2024), these could be applied to airline pilot training to improve data quality and facilitate oversight in order to maintain a fairer system.

The Regulators

The integration of AI into airline pilot training will be significantly shaped by regulatory frameworks. Regulators face a challenging task, as the field of AI is rapidly evolving. They must balance harnessing the benefits of AI with protecting the safety of those subject to their regulations, while also ensuring that ethical considerations are addressed.

Overall, the FAA's roadmap (FAA, 2024) seems to be more safety focused whereas EASA's roadmap (EASA, 2023), in addition to safety, explores ways of increasing investment and accelerating development. This may not completely reflect the current or upcoming regulations and their approaches as they may have different areas of responsibility and authority. The FAA is tied to a single country whereas EASA regulates the numerous member states of the EU.

Both roadmaps (FAA, 2024; EASA, 2023) show that these regulators are communicating with stakeholders, that they want to facilitate the implementation of AI in the aviation sector to benefit from it but recognise the risks and ethical considerations involved, although the ethical considerations are only explicitly stated by EASA.

Studies

The four studies included in this paper cover a wide array of training aspects. The study by Shaker & Al-Alawi (2023) provided a broad assessment of AI implementation in airline pilot training. Their findings highlighted the need for more personalised training (micro-adaptive learning) which could be achieved using AI due to its ability to process vast amounts of data.

The next study (Huang et al., 2019) was able to forecast flight school students learning curves. Although there is room for improvement, it did show that it was possible. With future refinements, if used by airlines, it could facilitate a more proactive approach to training by identifying and mitigating deficiencies earlier. This could increase training efficiency as deficiencies likely distract trainees from developing other competencies, fixing them early would allow more focus on other aspects of training. Furthermore, resources could be more accurately allocated where they are needed. This should be easily scalable, as it is mainly a digital solution only requiring data that already exists in training programs.

The following study (Dapica & Peinado, 2021), utilised AI to enhance the knowledge base for EBT training. By utilising AI, air safety report databases could be analysed thoroughly, providing a comprehensive view of competency deficiencies in the real world. By linking these competency deficiencies with competencies trained in training scenarios, training programs could be adapted in a timely and accurate manner. The process used, along with the resulting ontology, can serve as valuable tools to enhance an airline's EBT program by improving its thoroughness and allowing rapid adaptability. Once established, this should be relatively easy to scale, as it mainly requires data from sources that already exist. This would allow large organisations to be adaptive to new, emerging threats.

Finally, the fourth study (Ke et al., 2023) explored the use of AI in the form of ML to identify differences between expert pilots and novice pilots. If done on a larger scale, a comprehensive map of behaviours, movements, techniques as well as other characteristics could be created. Categorising based on core competencies in accordance with the EBT methodology rather than merely "expert" or "novice" could provide even more insight.

Identifying these types of differences could provide a basis for training. Being able to identify different, potentially subtle behaviours could help pilots find the root causes for why they cannot perform better in certain areas. The ability to identify these differences could also pave the way for training without an instructor, allowing pilots to train with VR headsets to complement their existing training.

Companies

Four companies were identified to be utilising AI to improve pilot training: CAE (n.d.-a), APC (n.d.), Hinfact (n.d.-b) and LAT (n.d.). CAE (n.d.-a) and Hinfact (n.d.-b) share a similar approach, offering comprehensive, systematically integrated end-to-end solutions that encompass the full training system. In contrast, APC (n.d.) and LAT (n.d.) focus on more targeted, specialised applications of AI, such as recruitment processes, scenario-based training, and specific simulator enhancements. This diversity in approaches highlights the broad potential for AI integration across various aspects of airline pilot training, from holistic management systems to focused tools addressing particular industry needs.

Both CAE (n.d.-a) and LAT (2024) incorporate VR into their projects. AI's capacity to process large volumes of data enables the creation of highly personalized training experiences. Digital environments can be rapidly, automatically, and extensively modified with minimal cost, allowing for greater flexibility. Combined with the immersive qualities of VR, these advancements have the potential to significantly enhance the affordability and scalability of pilot training.

The companies found to be utilising AI did not share much data on for example: measurable improvements in training efficiency, safety metrics or cost savings. They provided varying, but mostly limited information regarding how their products are practically used. CAE (n.d.-a) and Hinfact (n.d.-b) appeared to be more open about this in contrast to LAT (n.d.) and APC (n.d.). This is possibly due to financial reasons. Giving away too much information could compromise their profitability by giving away their unique advantage over competitors. This could also be because the latter two have not progressed as far in their development.

Despite this, the mere fact that large reputable companies are marketing and associating their brands with these solutions indicates that they have some degree of faith in its viability and effectiveness. If these solutions were to be ineffective, their reputations and brands could be negatively affected.

LAT Interview

The soft skills project by LAT, as detailed in chapter 3.7, utilises AI and XR to improve training. It achieves this mainly through increased immersion in contrast to traditional methods of teaching soft skills. In the future, collected data could be used to suggest training that aligns with the trainee's competency goals. Although not directly related, this technology could be adapted and utilised for airline pilot training. For example, pilots could train alone in VR cockpits with an AI driven captain or first officer. This could for example, help pilots practice how to utilise their colleagues better by practising their problem solving and decision-making process while asking their AI colleague open questions. The AI in the soft skill training project is also capable of checking whether certain recommended steps are taken. This could potentially be applied to flight deck procedures such as normal and emergency procedures or memory items to provide feedback. This could allow pilots to practise as needed at home to

complement their existing training programs. This extra practise could lead to a reduced need for costly full flight simulators, saving money in the long run.

Foyisi also mentioned the D-CEET project being worked on within LAT. This could potentially provide evacuation training for pilots as current training often ends with the evacuation checklist being completed in the flightdeck.

4.1 Limitations

Due to the broad scope of this study, it was not feasible to use a structured search for all segments. This limitation impacts the replicability of some segments of the research. Addressing the research question required a comprehensive understanding of the topic, which necessitated multiple searches and resulted in a substantial number of articles and sources. Given the volume of search results, some sources were primarily selected based on their prominence in search rankings. This approach may have introduced selection bias and led to an unbalanced representation of the available research.

Additionally, the field of AI is evolving rapidly, which means that some of the research cited in this paper may quickly become outdated. While efforts were made to include more recent sources where possible, some of the reviewed literature may not reflect the latest developments in AI.

As a result, the findings presented in this paper likely represent only a limited portion of the existing research. Future studies with a more structured methodology or a broader scope could provide a more comprehensive perspective and help validate or expand upon these findings. In hindsight, due to the limited available research, a broader approach, looking into the use of AI in airline training in general could have provided a better result.

5. Conclusion

Overall, there is limited research available on the use of AI in airline pilot training and limited information on its use by companies. This is likely because much of this research is conducted in-house by airlines and training providers. Airlines are likely reluctant to share their findings, as they have probably invested significant resources and seek to maintain a competitive edge. Similarly, training providers risk losing their competitive advantage if their research is shared, which could negatively impact profitability. Nevertheless, the available research demonstrates a broad range of potential implementations from forecasting pilot learning curves to identifying expert behaviours. The same can be said for its use by companies in pilot training.

APC (n.d.) and LAT (n.d.) utilized LLMs to create tools/virtual assistants in the form of chatbots, LAT (2024) also used LLMs to simulate human behaviour in their soft skills training project. In contrast to this a lot of research and companies used or explored AI in the form of analytical AI (CAE, n.d.-a; APC, n.d.; Hinfact, n.d.-b; Shaker & Al-Alawi, 2023; Huang et al., 2019; Dapica & Peinado, 2021; Ke et al., 2023). Both approaches have a potential to enhance training. The former could be a lower risk application as it merely provides an additional situation or scenario for trainees to add to their experience whereas the analytical approach directly addresses the trainee's performance.

Although there is a huge potential for AI use in airline pilot training, there are a few issues that must be considered. Deep learning, black box AI, algorithms lack transparency (Kosinski, 2024), meaning even their creators are unaware of how they work. This is problematic when used in safety critical industries. When the “thought process” of the algorithm cannot be verified, it would have to be trusted blindly. When things go wrong, liability becomes an issue.

Reliability is another problem with certain algorithms, due to the way they make inferences, they may hallucinate, generating incorrect information (Feuerriegel et al., 2023). AI models capable of hallucinating need to be carefully incorporated into training systems to ensure that incorrect outputs cannot cause unacceptable harm through thorough risk assessments.

Finally, ethical implications must be considered. The data and algorithms used could perpetuate existing biases or inadvertently affect individuals unfairly (Wajid et al., 2024). Regulators, researchers, airlines and training providers need to ensure appropriate safeguards are in place to ensure its use is ethical.

Future research could explore AI implementation in general education and discuss the potential to transfer some of this to airline pilot training. As mentioned in the introduction, there has been a significant increase in research on AI in education. Potential mitigations for some of the problems associated with AI, raised earlier in this paper, could also be assessed in the context of training.

6. References

References with an * represent the articles presented under 3.5.

The Airline Pilot Club. (n.d.). *The Airline Pilot Club*. <https://theairlinepilotclub.com/> [Accessed 15th May 2025]

The Airline Pilot Club. (2025, January 8). *APC Welcomes New CEO, Unveils 2025 Expansion for AI Platform Amelia Airline Pilot Club*. <https://theairlinepilotclub.com/candidates/news-events/apc-new-ceo-2025-amelia-expansion>

Amelia. (n.d.). *About Us*. <https://amelia.ai/about/> [Accessed 15th May 2025]

Barata, J.M.M. and Neves, F.M.S.P. (2017). *The History of Aviation Education and Training*. Open Journal of Applied Sciences, 7, 196-205. <https://doi.org/10.4236/ojapps.2017.74017>

Boeing. (2024). *Pilot and Technician Outlook 2024-2043*. <https://www.boeing.com/Commercial/market/pilot-technician-outlook#overview>

CAE. (n.d.-a). *CAE Trax Academy*. <https://www.cae.com/defense-security/what-we-do/training-systems/cae-trax/> [Accessed 15th May 2025]

CAE. (n.d.-b). *Our business*. <https://www.cae.com/about-cae/our-business/> [Accessed 15th May 2025]

- CAE. (n.d.-c). *CAE TraxAcademy*.
https://www.cae.com/media/documents/Defence_Security/CAE_Trax_Academy.pdf
 [Accessed 15th May 2025]
- Carbonaro, A. (2022). *Knowledge-based system as a context-aware approach for the Internet of medical connected objects*. Elsevier EBooks, 117–127.
<https://doi.org/10.1016/b978-0-32-391773-5.00012-1>
- Cooper, S. B., & J Van Leeuwen. (2016). *Alan Turing : his work and impact / Alan Turing: his work and impact*. Ha'erbin Gong Ye Da Xue Chu Ban She. (p. 481)
- *Dapica, R., & Peinado, F. (2021). *Towards a Semantic Knowledge Base for Competency-Based Training of Airline Pilots*. *Procedia Computer Science*, 192, 1208–1217.
<https://doi.org/10.1016/j.procs.2021.08.124>
- Durkin, J. (1990). *Research Review: Application of Expert Systems in the Sciences*.
<https://kb.osu.edu/server/api/core/bitstreams/98710464-d567-5dac-a77f-53e03a1a6ec9/content>
- European Commission. (2024). *AI Act*. <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>
- European Union Aviation Safety Agency. (2023, May 10). *EASA Artificial Intelligence Roadmap 2.0 - A human-centric approach to AI in aviation*.
<https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20>
- Federal Aviation Administration. (2024, July 23). *Roadmap for Artificial Intelligence Safety Assurance*. <https://www.faa.gov/media/82891>
- Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2023). *Generative AI*. *Business & Information Systems Engineering*, 66(1), 111–126. <https://doi.org/10.1007/s12599-023-00834-7>
- Flamm, K. (2018). *Measuring Moore's Law: Evidence from Price, Cost, and Quality Indexes*. NBER Working Paper No. 24553. National Bureau of Economic Research.
https://www.nber.org/system/files/working_papers/w24553/w24553.pdf
- Francesconi, E. (2022). *The winter, the summer and the summer dream of artificial intelligence in law*. *Artificial Intelligence and Law*. <https://doi.org/10.1007/s10506-022-09309-8>
- Goldman Sachs. (2023, April 5). *Generative AI could raise global GDP by 7%*.
<https://www.goldmansachs.com/insights/articles/generative-ai-could-raise-global-gdp-by-7-percent.html>
- Google DeepMind. (n.d.). AlphaGo. <https://deepmind.google/research/projects/alphago/>
- Haenlein, M., & Kaplan, A. (2019). *A brief history of artificial intelligence: On the past, present, and future of artificial intelligence*. *California Management Review*, 61(4), 5–14. Sage Journals. <https://doi.org/10.1177/0008125619864925>
- Hinfect. (n.d.-a). *About Us*. <https://hinfect.com/about-us/> [Accessed 18th May 2025]
- Hinfect. (n.d.-b). *Automated & Adapted Aviation Training Management System*.
<https://hinfect.com/automated-and-adapted-aviation-tms/> [Accessed 18th May 2025]
- Hinfect. (n.d.-c). *Training Management System*. <https://hinfect.com/atms-hinfect/> [Accessed 18th May 2025]
- *Huang, C.-L., Huang, C.-C., & Chiu, P.-H. (2019). *The Intelligent Forecasting Approach of Learning-curve in Fits Training of Student Pilot*. *International Journal of*

- Organizational Innovation (Online), 12(1), 205–214.
<https://www.proquest.com/docview/2251985985?sourcetype=Scholarly%20Journals>
- Huyck, C. R., & Passmore, P. J. (2021). *The Hebb Synapse Before Hebb: Theories of Synaptic Function in Learning and Memory*. *Frontiers in Behavioral Neuroscience*, 15, 732195. <https://doi.org/10.3389/fnbeh.2021.732195>
- IBM. (n.d.). *Deep Blue*. <https://www.ibm.com/history/deep-blue>
- IBM. (2021, October 6). *What is a Neural Network?*
<https://www.ibm.com/think/topics/neural-networks>
- IBM. (2023, December 12). *Support Vector Machine*.
<https://www.ibm.com/think/topics/support-vector-machine>
- International Air Transport Association. (2014). *Data Report for Evidence-Based Training Edition 1st*.
<https://www.iata.org/contentassets/c0f61fc821dc4f62bb6441d7abedb076/data-report-for-evidence-basted-training-ed20one.pdf>
- International Air Transport Association. (2024a, August 19). *Data at the heart of AI success. Airlines*. <https://airlines.iata.org/2024/08/19/data-heart-ai-success> [accessed on 20th of February 2025]
- International Air Transport Association. (2024b). *Global Outlook for Air Transport Deep Change*. <https://www.iata.org/en/iata-repository/publications/economic-reports/global-outlook-for-air-transport-june-2024-report/>
- International Civil Aviation Organization. (2013). *Manual of Evidence-based Training (Doc 9995) (1st ed.)* <https://skybrary.aero/sites/default/files/bookshelf/3177.pdf>
- Jones, C. R., & Bergen, B. K. (2025). *Large Language Models Pass the Turing Test*.
<https://doi.org/10.48550/arXiv.2503.23674>
- *Ke, L., Zhang, G., He, J., Li, Y., Li, Y., Liu, X., & Fang, P. (2023). *Pilot Selection in the Era of Virtual Reality: Algorithms for Accurate and Interpretable Machine Learning Models*. *Aerospace*, 10(5), 394. <https://doi.org/10.3390/aerospace10050394>
- Kinney, J. B., & Atwal, G. S. (2014). *Equitability, mutual information, and the maximal information coefficient*. *Proceedings of the National Academy of Sciences*, 111(9), 3354–3359. <https://doi.org/10.1073/pnas.1309933111>
- Kaplan, A., & Haenlein, M. (2019). *Siri, Siri, in My hand: Who's the Fairest in the land? on the interpretations, illustrations, and Implications of Artificial Intelligence*. *Business Horizons*, 62(1), 15–25. <https://doi.org/10.1016/j.bushor.2018.08.004>
- Kosinski, M. (2024, October 29). *Black box AI*. IBM.
<https://www.ibm.com/think/topics/black-box-ai>
- Lighthill, J. (1973). *Artificial intelligence: A general survey (Report to the Science Research Council)*. Science Research Council. https://rodsmitn.nz/wp-content/uploads/Lighthill_1973_Report.pdf
- Lufthansa Aviation Training. (n.d.). *Data-based training and artificial intelligence*.
<https://www.lufthansa-aviation-training.com/data-based-training-and-artificial-intelligence>
- Lufthansa Aviation Training. (2024, February 27) *Lufthansa Aviation Training (LAT) and 3Spin Learning Develop AI-Supported Crew Training With Apple Vision Pro*.
<https://www.lufthansa-aviation-training.com/-/lat-und-3spintraining-develop-ai-supported-crew-training>

- Marr, B. (2023, May 19). *A short history of ChatGPT: How we got to where we are today*. Forbes. <https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/>
- Moore, G. E. (1965). *Cramming more components onto integrated circuits*. Electronics, 38(8), 114–117. https://hasler.ece.gatech.edu/Published_papers/Technology_overview/gordon_moore_1965_article.pdf
- Minsky, M., & Papert, S. (1969). *Perceptrons: An Introduction to Computational Geometry*. MIT Press.
- Myers, P. L., Starr, A. W., & Mullins, K. (2018). *Flight Simulator Fidelity, Training Transfer, and the Role of Instructors in Optimizing Learning*. International Journal of Aviation, Aeronautics, and Aerospace, 5(1). <https://doi.org/10.15394/ijaaa.2018.1203>
- Patil, N. H., Patel, S. H., & S.D. Lawand. (2023). *Research Paper On Artificial Intelligence And It's Applications*. Journal of Advanced Zoology, 44(S8). <https://doi.org/10.53555/jaz.v44is8.3544>
- PWC. (2021, June 4). *How Virtual Reality Is Redefining Soft Skills Training*. <https://www.pwc.com/us/en/tech-effect/emerging-tech/virtual-reality-study.html>
- Raman, R., Kowalski, R., Krishnashree Achuthan, Iyer, A., & Prema Nedungadi. (2025). *Navigating artificial general intelligence development: societal, technological, ethical, and brain-inspired pathways*. Scientific Reports, 15(1). <https://doi.org/10.1038/s41598-025-92190-7>
- Roza Sh. Akhmadieva, Kalmazova, N. A., Belova, T., Alexey Prokopyev, Molodozhnikova, N. M., & Spichak, V. Y. (2024). *Research trends in the use of artificial intelligence in higher education*. Frontiers in Education, 9. <https://doi.org/10.3389/educ.2024.1438715>
- Sarker, I. H. (2021). *Machine Learning: Algorithms, Real-World Applications and Research Directions*. SN Computer Science, 2(3), 1–21. Springer. <https://doi.org/10.1007/s42979-021-00592-x>
- Sarker, I. H. (2022). *AI-Based Modeling: Techniques, Applications and Research Issues Towards Automation, Intelligent and Smart Systems*. SN Computer Science, 3(2). Springer. <https://doi.org/10.1007/s42979-022-01043-x>
- *Shaker, M. H., & Al-Alawi, A. I. (2023). *Application of Big Data and Artificial Intelligence in Pilot Training: A Systematic Literature Review*. 2023 International Conference on Cyber Management and Engineering (CyMaEn). <https://doi.org/10.1109/cymaen57228.2023.10050972>
- Smith, C., McGuire, B., Huang, T., & Yang, G. (2006). *The History of Artificial Intelligence*. University of Washington. <https://courses.cs.washington.edu/courses/csep590/06au/projects/history-ai.pdf>
- Tao, W., Zeng, H., & Rong, X. (2014). *Research on Design Assurance System with the Development of CAAC's Design Organization Management*. Procedia Engineering, 80, 224–230. <https://doi.org/10.1016/j.proeng.2014.09.082>
- Trezza, A., Visibelli, A., Roncaglia, B., Ottavia Spiga, & Santucci, A. (2024). *Unsupervised Learning in Precision Medicine: Unlocking Personalized Healthcare through AI*. Applied Sciences, 14(20), 9305–9305 <https://doi.org/10.3390/app14209305>

- Turing, A. M. (1938). *On Computable Numbers, with an Application to the Entscheidungsproblem. A Correction*. Proceedings of the London Mathematical Society, s2-43(1), 544–546. <https://doi.org/10.1112/plms/s2-43.6.544>
- Velarde, G. (2019). *Artificial Intelligence and its Impact on the Fourth Industrial Revolution: A Review*. International Journal of Artificial Intelligence & Applications, 10(6), 41–48. <https://doi.org/10.5121/ijaia.2019.10604>
- Wajid, A., None Musawer Hakimi, & Amir, N. (2024). *A Comprehensive Review of Bias in AI Algorithms*. Nusantara Hasana Journal, 3(8), 1–11. <https://doi.org/10.59003/nhj.v3i8.1052>
- Zadeh, L. A. (2008). *Is there a need for fuzzy logic?* Information Sciences. 178(13), 2751–2779. <https://doi.org/10.1016/j.ins.2008.02.012>
- Zafar, M. R., & Khan, N. (2021). *Deterministic Local Interpretable Model-Agnostic Explanations for Stable Explainability*. Machine Learning and Knowledge Extraction, 3(3), 525–541. <https://doi.org/10.3390/make3030027>
- Zhu, F., Xu, P., & Zong, J. (2023). *Moore's Law: The potential, limits, and breakthroughs*. Applied and Computational Engineering, 10(1), 307–315. <https://doi.org/10.54254/2755-2721/10/20230038>

