The Growing Role of Artificial Intelligence (AI) in Aviation Safety and the Necessity to Create Strong Protocols in the Context of Current Global Conditions and with the Advent of Autonomous Flight

ISASI Rudolph Kapustin Memorial Scholarship 2023

Sebastian Torres

University of Southern California

Viterbi School of Engineering, Aviation Safety & Security Program

B.S. / M.S. in Aerospace Engineering

Supervisors: Thomas Anthony and Daniel Scalese

Introduction

With the emergence of autonomous self-driving vehicles, and the global Artificial Intelligence (AI) and Machine Learning market in aviation projected to reach $3.4 billion by 2027 (Airline Industry Information), it becomes increasingly imperative to examine autonomous systems in the context of Aircraft Accident Investigation (AAI). The aviation industry is no stranger to the implementation of AI technologies including, but not limited to, predictive airplane maintenance, airborne collision avoidance systems (ACAS), flight path optimizations, and more recently, the research and development of fully autonomous flight. This paper focuses on AI's benefits and why it will continue to expand with aviation, establish the security and safety concerns AI poses, and make suggestions on how investigators should navigate the creation of AI certifications and protocols in light of current global economic conditions.

Brief Overview of Successful AI Implementation

To begin, a quick summary of current usage of AI in aviation will be presented. Notable examples of autonomous flight include Garmin's AUTONOMÍ suite; automatic ground collision avoidance system (Auto GCAS) developed by Lockheed Martin, NASA, and the Air Force; and Aurora Flight Sciences's uncrewed, optionally piloted aircraft. In particular, the Autoland in AUTONOMÍ was awarded the Collier trophy in 2020 (Garmin). Auto GCAS also has been responsible for nine F-16s saved since its adoption in the U.S. Air Force in 2014 (Lockheed Martin).

Other applications of artificial intelligence in aviation come in the forms of pilot aid and efficiency optimization. These include decision support systems and intelligent crew interfaces for pilots to lessen emotional and sensory loads, reduce reaction time, active monitoring of physical and psychological conditions, and perform difficult tasks like "object recognition and tracking [and] synthesis of safe flight paths" (Kulida and Lebedev). As for efficiency optimization, AI can improve air traffic management using optimization algorithms to process large data sets, generate optimal flight routes, and create landing sequences and timings.

While these implementations provide a promising insight into what AI can accomplish in aviation, there are several points of safety and security concerns investigators must consider.

Security Concerns Regarding Data

One concern regarding the security of AI lies in the data used to train the machine learning algorithms. Data, like aircraft parts, must be tested and certified. Cheaper parts or data may be counterfeit or misrepresented.

Care needs to be taken to use data as a tool, not the ultimate deciding factor, in machine learning algorithms. Too much emphasis on data can result in overfitting, the phenomenon where the algorithm creates incredibly accurate results based on one set of data but fails to find further meaningful predictions with new data. Another issue can result where critical details and outliers that represent significant but rare occurrences are ignored to achieve greater correlations. Investigators cannot ignore the hazards created by incomplete, synthetic, or biased data. When training machine learning algorithms, it is imperative that the principle of "correlation does not equal causation" be employed.

Another point of concern is the inherent randomness and non-deterministic nature of the algorithms themselves. While algorithms are being trained, the same output cannot necessarily be guaranteed each time for a given input (EASA). Investigators must know how to work with this uncertainty.

Economic Disincentives

Current global market conditions have also exacerbated the current issues surrounding AI. Prior to the events of Silicon Valley Bank (SVB) and Credit Suisse in Spring 2023, the financial markets experienced low interest rates after rebounding from the onset of COVID-19, which resulted in investors seeking greater yields. This incentivized investors to make highly speculative investments and raise the valuations of certain technologies, which included following the tech sector's trend of AI development. The focus was not to evaluate the efficacy or profitability of the technology in question, but rather use it as a vehicle by which to move large sums of money. This disincentivizes performing due diligence in the creation of AI systems, which directly opposes the deterministic assurance investigators seek (Nield). Those involved in the research of implementing AI in aviation-based systems may experience pressures to deliver successful products, which can result in models that are too optimistic and final products plagued with hazards and safety concerns. This dynamic tension between production and protection has been at the heart of aviation safety for decades.

Overstatements of AI Capabilities

Another issue that results from economic disincentivization and overoptimism in AI is overstating its capabilities for better marketing. Liza Dixon, a doctoral candidate at Universität Ulm researching human-machine interaction in automated driving, coined the term "autonowashing" (Dixon) to describe this very gap in stated capability and actual results with AI. Tesla's misuse of the term "autopilot" and the death of Elaine Herzberg from an autonomous Uber car are two illustrative examples.

Tesla is known for marketing their vehicles as having “autopilot and full self-driving capability.” Confusion quickly arises with combining these two terms. The suite has three tiers, each with increasing assistance features: autopilot, enhanced autopilot, and full self-driving capability. Despite what the highest tier’s name suggests, none of these actually make the vehicle autonomous and thus still require driver supervision at all times (Tesla). The average consumer can easily assume that a Tesla equipped with “autopilot” can drive on its own. A misunderstanding of or overreliance on these systems endangers other drivers on the road and can lead to accidents, or worse, fatalities.

Such a fatal incident occurred in March 2018, when a Tesla Model X collided with a crash attenuator on the US Highway 101. This investigation was notably difficult, as Elon Musk hung up on a phone call with the NTSB Chair (Ars Technica). Tesla would eventually be removed as a party in the investigation after prematurely releasing confidential investigative information. The NTSB found the probable cause of the accident to be a combination of the autopilot system’s limitations, the driver’s overreliance on the system, and his lack of response due to being distracted by a mobile game. As former National Transportation Safety Board (NTSB) Chair Robert L. Sumwalt articulated, “If you own a car with partial automation, you do not own a self-driving car” (NTSB). This delineation can and should be applied to every sector of transportation simply by replacing “car,” with “plane”, “bus,” etc.

The Uber incident in Tempe, Arizona more clearly shows an algorithmic failure. It was determined that the modified Volvo XC90’s LIDAR and radar pinged the jaywalking pedestrian 6 seconds prior to impact while it was traveling at 43 mph. However, as the car continued to approach the pedestrian, its systems would misclassify the person as an unknown object, vehicle, and bicycle. The system also had a “prioritization schema” (NTSB), which placed emphasis on recent observations. At that moment, the car was predicting a different trajectory than if it were correctly identifying the pedestrian. As a result, the emergency detection only provided the backup human operator with ~1.3 seconds of time to react. This accident proved that logical inconsistencies in AI need to be thoroughly explored, proper safety systems must be in place, and algorithms must be adequately equipped to account for unexpected conditions like jaywalking (Nield).

While both instances relate to autonomous driving, investigators should study these cases to develop rigorous protocols surrounding all applications of autonomy in aviation. No disservice should be done to the traveling public, and no official should have to face the difficulty of investigation removal due to a failure or refusal of a manufacturer to explain their software.

As demonstrated just in this paper alone, the term "artificial intelligence" has also become a blanket term to encompass everything from chat bots to automated flight path optimization systems being developed. The failure to deliver on promises and a continued dilution of the words "Artificial Intelligence" (Dixon) results in an attitude of distrust. In a sector of transportation where people already place their trust on several factors to take them safely to their destinations, including the manufacturers of the planes and engines, the pilots' expertise, and air traffic control, there must be trust in the technology of AI or else the public will not accept it. It is therefore the duty of air safety investigators and policymakers to develop detailed certification programs and licenses for AI implementation.

Pragmatic Action Forward and Valuation of Safety

The continued successful deployment of AI-based systems in aviation will rely on how much trust we put in the technology. That trust must be continuously earned through successful performance. In the context of autonomous vehicles, trust in automation is defined as: “the attitude of a user to be willing to be vulnerable to the actions of an automated system based on the expectation that it will perform a particular action important to the user, irrespective of the ability to monitor or to intervene” (Körber, 2018).

One great example of building trust in AI to learn from is Rolls Royce’s Aletheia Framework. They describe Aletheia as a “practice one-page toolkit that guides developers, executives and boards both prior to deploying an AI, and during its use” (Rolls-Royce). It states 32 principles covering areas of social impact, accuracy and trust, and governance to navigate the ethical intricacies of AI application. Interdisciplinary discussion about AI will drive safe innovation and realize its potential to improve society.

Designing technologies and systems to optimize aviation without exploring the opportunities AI has to offer would simply overlook its potential (Garcia et al.). Therefore, responsible actions must be taken to continually ensure the safety of passengers as technology moves forward.

Works Cited

"The Aletheia Framework® | Rolls-Royce" Available: https://www.rolls-royce.com/innovation/the-aletheia-framework.aspx.

"Artificial Intelligence Roadmap: A human-centric approach to AI in aviation," *EASA | European Union Aviation Safety Agency* Available: https://www.easa.europa.eu/sites/default/files/dfu/EASA-AI-Roadmap-v1.0.pdf.

"Automatic Ground Collision Avoidance System (auto GCAS)," *Lockheed Martin* Available: https://www.lockheedmartin.com/en-us/products/autogcas.html.

"Autonomí® Autonomous Safety-Enhancing Technologies," *Garmin Autonomí® | Autonomous Flight Solutions* Available: https://discover.garmin.com/en-US/autonomi/.

"Autopilot and full self-driving capability: Tesla Support ," *Tesla* Available: https://www.tesla.com/en_eu/support/autopilot.

"Board Meeting: Collision Between a Sport Utility Vehicle Operating With Partial Driving Automation and a Crash Attenuator - Opening Statement," *National Transportation Safety Board (NTSB)* Available: https://web.archive.org/web/20200303040151/ https://www.ntsb.gov/news/speeches/RSumwalt/Pages/sumwalt-20200225o.aspx.

"Collision Between Vehicle Controlled by Developmental Automated Driving System and Pedestrian," *National Transportation Safety Board (NTSB)* Available: https://www.ntsb.gov/investigations/AccidentReports/Reports/HAR1903.pdf.

Dixon, L., "Autonowashing: The Greenwashing of Vehicle Automation," *Transportation Research Interdisciplinary Perspectives*, vol. 5, 2020.

Farivar, C., "Elon Musk hung up on NTSB chief during call about Tesla Crash probe," *Ars Technica* Available: https://arstechnica.com/tech-policy/2018/05/elon-musk-hung-up-on

-ntsb-chief-during-call-about-tesla-crash-probe/.

Garcia, A. B., Babiceanu, R. F., and Seker, R., “Artificial Intelligence and machine learning approaches for aviation cybersecurity: An overview,” *2021 Integrated Communications Navigation and Surveillance Conference (ICNS)*, Apr. 2021.

Kulida, E., and Lebedev, V., “About the Use of Artificial Intelligence Methods in Aviation,” *2020 13th International Conference "Management of large-scale system development" (MLSD)*, Sep. 2020.

Körber, M., “Theoretical considerations and development of a questionnaire to measure trust in automation,” *Advances in Intelligent Systems and Computing*, Aug. 2018.

Nield, T., “Machine Learning Explainability,” *University of Southern California Viterbi School of Engineering*.

Nield, T., “Uber Tempe Incident,” *University of Southern California Viterbi School of Engineering*.

“Worldwide Artificial Intelligence in Aviation Industry is Expected to Reach USD 3.4bn by 2027, According to Research and Markets,” *Airline Industry Information*, 2022.