MERLIN

Innovate to Elevate

Automation's Role in the Ongoing Evolution of Aviation Safety

January 2025

Executive Summary

120 years after the first powered flight, aviation continues to be one of mankind's greatest achievements. In all aspects of modern aviation, including general aviation, commercial airliners, and even defense applications, the best flying happens when operators have the best technology, and today, aviation stands proudly as the safest transportation mode in our technologically infused world. Staying ahead of the technological curve, however, is no small task. Safe and secure flight operations are a journey, not a destination, and we must continually strive for excellence amidst an ever-evolving environment.

While technology continues to advance, the human brain's unparalleled ability to assess and interpret complex situations, and also make critical judgments, remains the cornerstone of safe flight operations, especially when hundreds of lives are at stake with each and every flight. At Merlin, we believe the next leap in aviation safety will come from pairing the strengths of existing human pilots with the benefits of advanced automation. Automation already plays a vital role in helping pilots' hands, eyes, and brains focus on the most critical tasks in the flight deck, but moving forward, it must evolve from a tool into a trusted partner that enhances decision-making, reduces workload, and increases performance.

Enabling this future will require deliberate, methodical steps that incrementally build the capability of emerging technologies, all while remaining grounded in proven safety practices and working collaboratively with industry and regulators. This paper discusses how we can take the first of these steps today, using automation technologies that already exist.

This paper presents three tangible ways existing automation technology can complement the skills of human pilots to enhance flight safety. First, it discusses how adopting a human-centered design philosophy could better manage flight crew workload, reducing incidents associated with operator task saturation or inattentiveness. Next, it discusses the use of consumer wearables to better understand pilot fatigue, which would enable bespoke and more accurate approaches to fatigue management. Finally, it explores how automation can increase safety by reducing miscommunications in pilot and air traffic controller radio transmissions.

These low-cost advancements come with a high benefit—enabling pilots to focus on their primary responsibilities to aviate, navigate, and communicate free of distractions during critical moments. Implementing these technologies requires collaborative effort between the public and private sectors, and Merlin looks forward to working with industry experts, pilots, air traffic controllers, and regulators to pave the way for higher levels of automation grounded in proven safety practices, for the benefit of all.

Table of Contents

MERLIN

HEADQUARTERS: 129 South Street, Boston, MA 02111

Introduction

Commercial aviation is firmly established as the safest mode of transportation available, achieving record levels of safety in 2023[1]. However, that same year also saw a record number of runway incursion incidents, a renewed focus on operator fatigue and mental health, and continued discourse regarding the importance of aircraft certification and quality management. These events remind us that the National Airspace System (NAS) is constantly being challenged in new and unforeseen ways, and that aviation safety will always be an ongoing and ever-evolving journey.

In commercial aviation, having two sets of eyes and hands in the flight deck has long been foundational to safety. There is simply no replacing the human brain when hundreds of lives are at stake with each and every flight. At Merlin, we believe the key to meeting modern flight safety challenges lies in pairing the unique strengths of human pilots with the benefits of automation. We see a future where pilots and automation form a more advanced and robust partnership—extending human capability with technology that does more and responds more accurately and quickly. Furthermore, defense applications will provide key opportunities for enabling this future, as the military has long been a proving ground for innovations that later transform the private sector. This is why Merlin is committed to pursuing dual-use applications that leverage the speed of advancements in military operations to inform a deliberate, incremental approach for commercial transport aircraft.

Since 2018, we have been developing technical solutions that augment the skills of a human pilot in a variety of civil and military applications. Automation, though not new to aviation, has advanced significantly and still represents an area of untapped potential. By embracing technological innovation with a pilot-first view, we believe the aviation community can unlock enhanced levels of safety and performance to address current and future challenges while simultaneously enabling new technologies and use cases.

The long-term vision of automation[2] and autonomy[3] in commercial aviation, combined with the many economic, labor, and technology changes that follow, are too vast for one whitepaper or company to address. For now, we want to start a dialogue with industry, regulators, and other stakeholders about how automation can better complement human expertise in commercial transport aircraft and increase safety in the near term.

[1]International Air Transport Association (IATA). (2024). *2023 Safest Year for Flying By Several Parameters*. IATA. https://www.iata.org/en/pressroom/2024-releases/2024-02-28-01/.

[2]Merlin appreciates that the words "automation" and "autonomy" are holistic terms that can have different meanings across different contexts and industries. With regards to aviation and Advanced Aerial Mobility specifically, Merlin encourages use of the terms and definitions established by ASTM International's AC-377 Working Group in their technical report "Autonomy Design and Operations in Aviation: Terminology and Requirements Framework" (DOI: 10.1520/TR1-EB). The document is specifically aimed at discussing the spectrum of automation, including when and how a system might be considered "automated" versus "autonomous", and serves as a valuable baseline the industry can adopt in trying to align terminology.

[3]See above.

MERLIN

HEADQUARTERS: 129 South Street, Boston, MA 02111

This paper outlines three critical safety imperatives that pose ongoing challenges to aviation safety and discusses how automation that more deeply considers the pilot can best address them:

1. **Better Workload Management:** A renewed focus on human-machine teaming can better inform automation systems design, ensuring that operators are neither task-saturated nor disengaged.
2. **Reduced Crew Fatigue:** Passive consumer wearables can lend insights into the relationship between stress, workload, and fatigue, leading to more robust fatigue mitigation tools.
3. **Clearer Radio Communications:** Natural Language Processing (NLP) technology can improve communication and coordination between pilots and air traffic controllers, reducing readback errors and miscommunications.

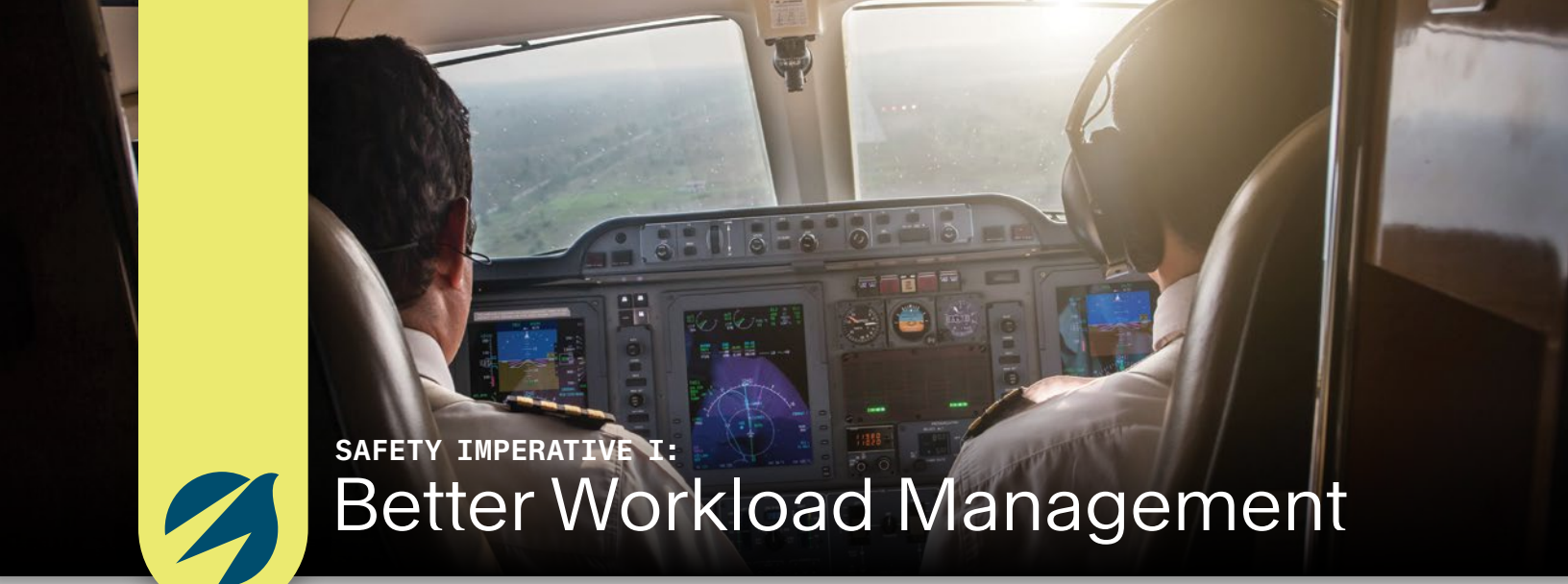
These safety imperatives are not exhaustive, but they can leverage current technology to further improve aviation safety in the near-term, while also laying the foundation for future advancements in safety-focused automation and autonomy. This approach aligns with recommendations from industry leaders and the airline pilot community, which emphasize the importance of integrating informational tools and systems under pilot oversight as a "lower-risk" pathway to introducing artificial intelligence (AI) in commercial aviation[4].

Finally, the paper concludes with near-term recommendations for regulators, policymakers, and industry to pursue in driving the integration of automation technologies and further enhancing safety. We will spell out recommendations to:

1. Create aviation communication datasets for NLP training
2. Quantify aviation communication performance metrics
3. Streamline a pathway for integrating AI technologies

One of our core beliefs at Merlin is advancing automation solutions in an incremental nature. We do not advocate for—nor do we envision—a world in which robots are flying commercial aircraft or controlling our skies. Rather, we believe in an incremental approach. Crawling before walking, and walking before running. By leveraging the recent proliferation of automation technologies in logical and incremental ways, the entire aviation community can take safety to new heights, making the next generation of flying the safest yet.

[4]In response to the U.S. Department of Transportation's Request for Information (RFI) regarding "Opportunities and Challenges of Artificial Intelligence (AI) in Transportation" (Docket No. DOT-OST-2024-0049), Air Line Pilots Association, International (ALPA) suggested that "new forms of informational development and guidance with pilot oversight and management may serve as a lower-risk approach to introducing AI into traditionally piloted airline aircraft."



SAFETY IMPERATIVE I:

Better Workload Management

Workload and task management have long been a challenge to aviation safety, as they directly relate to a pilot's ability to maintain situational awareness and make critical judgments. It's an extremely well documented problem, as as the National Aeronautics and Space Agency's (NASA's) Aviation Safety Reporting System (ASRS) regularly highlights "task saturation" as an ongoing safety risk[5]. Furthermore, its scope is wide reaching: problems with situational awareness were found to be the leading causal factor in a review of military aviation mishaps[6], and statistics reveal that approximately 85% of aviation accident and incident reports reference the loss of situational awareness[7].

In recent decades, advanced computing and automation solutions have proliferated across nearly every aspect of flight operations. Such systems are typically designed to aid pilots in high-pressure, high-workload scenarios where human performance is most vulnerable. However, when an automated system abruptly fails or acts unexpectedly, it can cause tremendous real-time stress, often leading to a loss of situational awareness and potentially an in-flight emergency. Therefore, a key opportunity to meet this safety challenge is designing automation tools that fail more gracefully, without unintended consequences and without putting undue burden on operators. At Merlin, we're putting this opportunity into practice by developing a design process that prioritizes human factors as a key input to aircraft systems design.

Historically, the approach to automation design has been to identify the tasks of a specific operation and then sort those tasks into those that can be easily automated and those that can't. Tasks that cannot or should not be assigned to machines are left for humans to perform. Since the determination of what is left over reflects what technology *cannot* easily do rather than what humans *can* do (or, more importantly, what humans *can do well*), the inevitable result is that human operators are faced with tasks that are too infrequent, too variable, or too expensive to automate. Further, humans are expected to "take over" with short notice should the system falter, which requires the system to provide rapidly digestible information to keep its operator fully apprised of what the system is doing, why it is doing it, and what it intends to do. This concept of "system handoff" remains a significant challenge for systems designers and operators, yet is critical to certification from a human factors perspective.

[5]Callback: From NASA's Aviation Safety Reporting System. (2019). *Distractions in the Cockpit*. National Aeronautics and Space Administration (NASA). https://asrs.arc.nasa.gov/docs/cb/cb_492.pdf

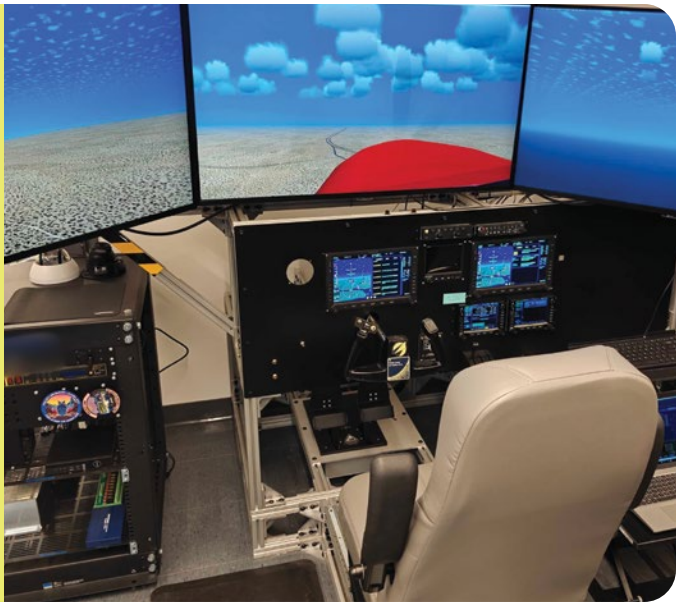
[6]Härtel, C. E. J., Smith, K. A., & Prince, C. (1989). Defining aircrew coordination: Searching mishaps for meaning.

[7]Skybrary. (2016). Situational awareness. https://skybrary.aero/articles/situational-awareness-oghfa-bn

Rethinking Automation Design to Improve Workload Management

Merlin is actively working with the United States military to create a more thoughtful automation design framework that naturally treats human factors as a primary consideration.

This work is merely the first step in our vision for evolving human factors design methodologies, but has already proven successful in more accurately designing for the strengths of the human machine team.



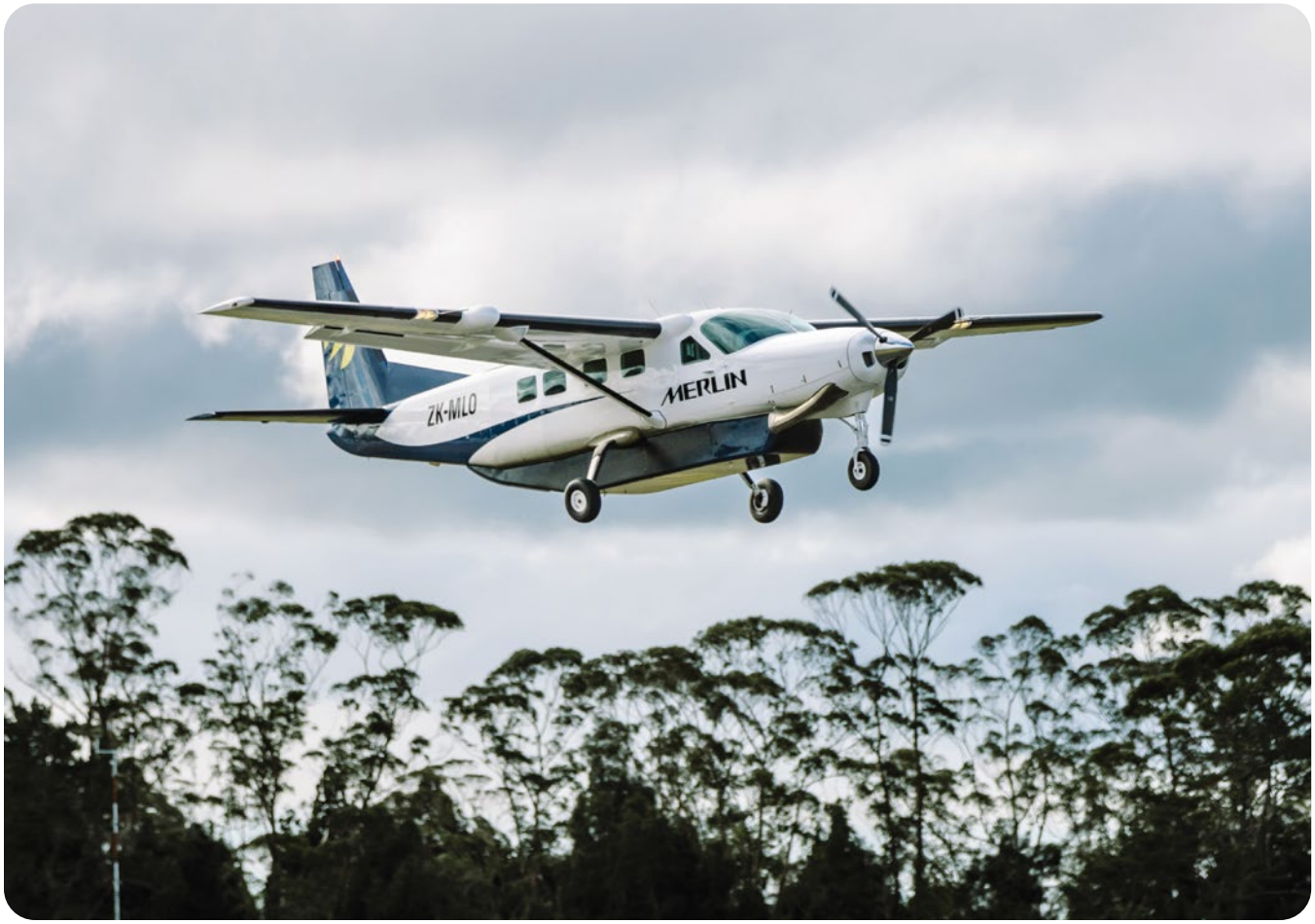
This design approach is known as the Leftover Principle[8], and in aviation often leads to a collection of illogical, incoherent tasks that are assigned to a human pilot, further leading to confusion, increased workload, and decreased situational awareness—all of which create opportunities for human error. The process of deciding what to automate based on how easily it can be achieved has also led to negative results: making the easy tasks easier while inadvertently making the harder tasks harder. For example, automated procedures during cruise (a phase of flight already associated with relatively low workload) have created opportunities for pilots to become disengaged or inattentive, leading to decreased situational awareness[9, 10].

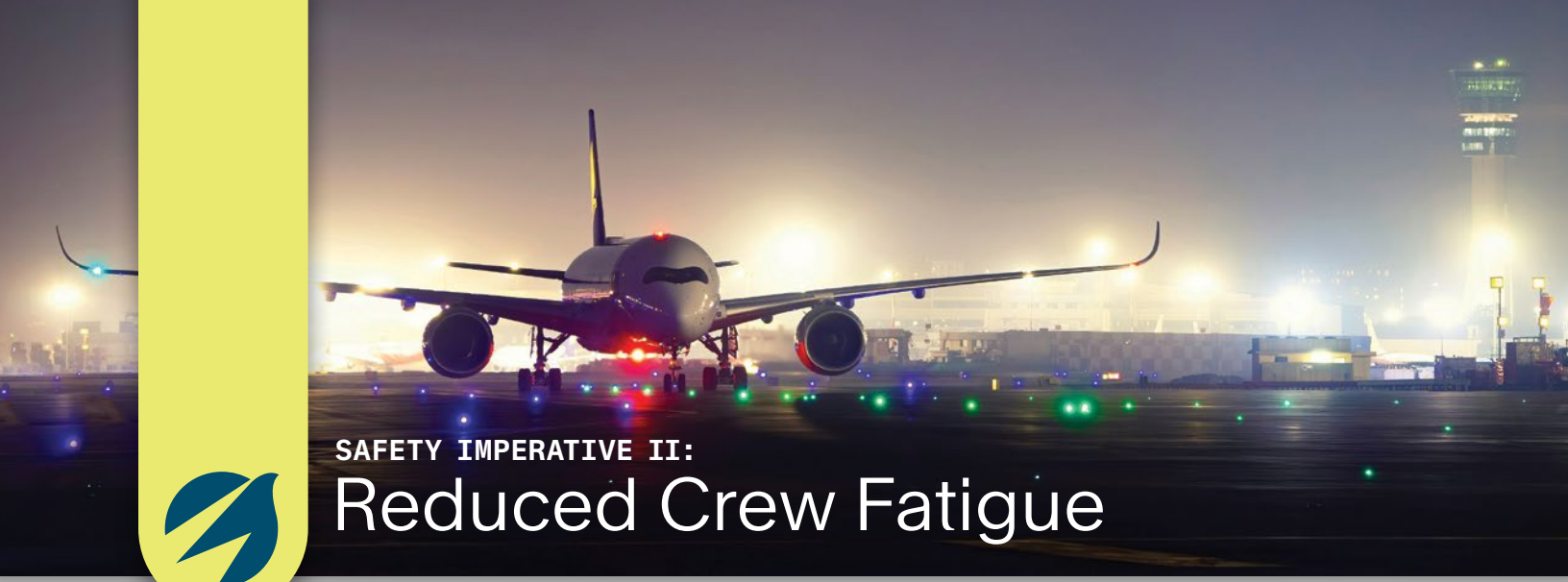
Automation development and integration should be a human-centered process that leverages the strengths of both the human and the system, referred to as "human-machine teaming". This approach ensures that the collection of tasks that ultimately rely on a human operator are coherent, logical, and balanced in the sense that they create neither increased workload nor opportunities for inattentiveness. Human-machine teaming requires thoroughly evaluating *when* automation is appropriate, *what* problems automation could help solve, and *how* to best integrate automation systems that complement human capacities.

Merlin is actively putting this human-centric design process to work today in partnership with the United States military as it works to integrate automation solutions into large transport category aircraft. Early in the design process, our engineers work closely with flight crew to analyze severe workload conditions and systematically deconstruct all related crew actions to identify the highest impact tasks and contextual factors. With this information, we can then design systems to best assist the pilot in areas where automation is most beneficial.

[8]Bainbridge, Lisanne. (1983). *Ironies of automation. Automatica.* 1983; Vol 19, Issue 6: 775-779.
[9]BBC. (2013). *'More than half' of pilots have slept while flying.* https://www.bbc.com/news/uk-24296544
[10]Werfelman, Linda. (2009). *Asleep at the wheel. Flight Safety Foundation.* https://flightsafety.org/asw/sept09/asw_sept09_p24-28.pdf

More broadly, our approach is to create a more efficient design framework that naturally treats human factors as a primary consideration. For every task assistance challenge, there typically exists a range of automation strategies with varying levels of invasiveness, complexity, and criticality. By developing new and more robust methods of quantifying workload, fatigue, and alertness, we can more accurately simulate and model the impacts of different automation solutions across a range of different scenarios. This model then allows us to rapidly identify the best automation solution for a given situation, and also obtain validation metrics to confirm that safety enhancements are successful in the finished product.





SAFETY IMPERATIVE II:

Reduced Crew Fatigue

Fatigue has long been a challenge to aviation safety, as it is linked to increased variability and unreliability in performance. Pilots are particularly vulnerable to fatigue due to long working hours, off-nominal report times, and circadian rhythm disruption, all of which cause slower reaction times, decreased cognitive function, and increased error rates[11].

In recent years, record passenger demand and persistent crew shortages have intensified pilot workload, causing renewed calls for action regarding crew fatigue. Air traffic controllers, who are also critical to aviation safety, face similar challenges. A 2023 report by the U.S. Department of Transportation revealed that 77% of critical air traffic control facilities operate below the FAA's 85% staffing threshold[12], often forcing controllers into six-day workweeks[13]. These realities on the ground and in the skies make fatigue a systemic issue across commercial transport aviation.

The connection between fatigue and aviation safety is undeniable. NASA's ASRS cites fatigue as a factor in 52,000 incidents—21% of all reports—underscoring its significant role in operational risk[14]. Additionally, sleep scientists agree that fatigue is one of the largest preventable causes of accidents across all transportation sectors[15].

Addressing this challenge requires us to do more than ensure operators are well-rested. Instead, we need to consider how newer, more robust automation tools can unlock a better overall approach to fatigue management and create a better flying experience for pilots.

The aviation industry currently uses Fatigue Risk Management Systems (FRMS) to model human performance and develop pilot schedules that best manage fatigue. However, FRMS typically use generalized parameters that don't consider an individual's biology or how well rested that individual actually is when reporting for duty[16]. A similar approach is also evident in the regulations that govern flight

[11]Federal Aviation Administration (FAA). (2010). *Advisory Circular 120 - 100.* U.S. Department of Transportation. https://www.faa.gov/documentLibrary/media/Advisory_Circular/AC%20120-100.pdf

[12]Office of Inspector General. (2023). *FAA Faces Controller Staffing Challenges as Air Traffic Operations Return to Pre-Pandemic Levels at Critical Facilities.* U.S. Department of Transportation. https://www.oig.dot.gov/sites/default/files/FAA%20Controller%20Staffing%20and%20Training%20at%20Critical%20Facilities%20Final%20Report-06-21-23.pdf

[13]National Air Traffic Controllers Association (NATCA). (2024). *NATCA Calls on FAA to Collaborate on Air Traffic Controller Fatigue.* NATCA. https://www.natca.org/2024/04/19/natca-calls-on-faa-to-collaborate-on-air-traffic-controller-fatigue/

[14]National Transportation Safety Board. (2003). *Commercial transportation operator fatigue management.* NTSB.

[15]Akerstedt, Torbjorn. (2000). *Consensus Statement: Fatigue and accidents in transport operations.* Journal of Sleep Research, Vol. 9 Issue 4.

[16]Federal Aviation Administration (FAA). (2013). *Advisory Circular 120 - 103A.* U.S. Department of Transportation. https://www.faa.gov/documentlibrary/media/advisory_circular/ac_120-103a.pdf

MERLIN

HEADQUARTERS: 129 South Street, Boston, MA 02111

crew duty limits, where basic tables are used to relate a pilot's report time with the maximum duty time that pilot may perform[17].

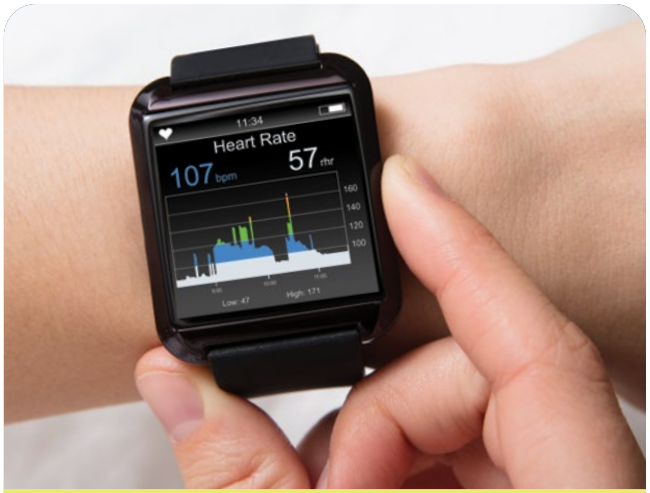
These 'one-size-fits-all' approaches to fatigue management leave gaps that automation technologies can help fill.

Right now, Merlin is actively testing with passive consumer wearables to help us better understand and mitigate fatigue. Data from wearables more accurately conveys measures of sleep quality, quantity, and recovery parameters, and in the aviation domain could alleviate the need for many of the assumptions made in current FRMS. Furthermore, wearables provide a non-invasive method of measuring physiological parameters during flight operations that could lend further insights into stress, workload, and other factors.

Our research, conducted in partnership with commercial operators, is not focused on determining a pilot's immediate fitness for duty. Instead, we're leveraging biometric data to better understand how trends in rest, sleep, and workload impact pilot performance. Early feedback from our partners has been encouraging and is already helping us understand how pilot stress and workload evolve over time.

This work isn't being done in isolation. Other safety-critical industries are already demonstrating how data-driven tools can enhance safety and operational efficiency, without getting in the way of the job at hand.

For first-responders, automation tools have yielded measurable reductions in overexertion injuries, vehicle accidents, and burnout in trial programs across the country[18, 19]. And in high-pressure healthcare settings, doctors and nurses are wearing passive devices that track heart rate, heart rate variability, and stress to improve performance and attention during long shifts[20, 21].



Leveraging Wearable Technology to Reduce Crew Fatigue

A better understanding of pilot fatigue, and its relationship to stress and workload, will be a critical requirement for advancing automation over the long term.

Merlin is partnering with commercial operators to better understand how consumer wearables can offer insights regarding pilot stress, workload, and fatigue.

[17]14 CFR Part 117, Tables A and B.

[18]Patel, V., Chesmore, A., Legner, C.M. and Pandey, S. (2022). T*rends in Workplace Wearable Technologies and Connected-Worker Solutions for Next-Generation Occupational Safety, Health, and Productivity*. Adv. Intell. Syst., 4. https://doi.org/10.1002/aisy.202100099

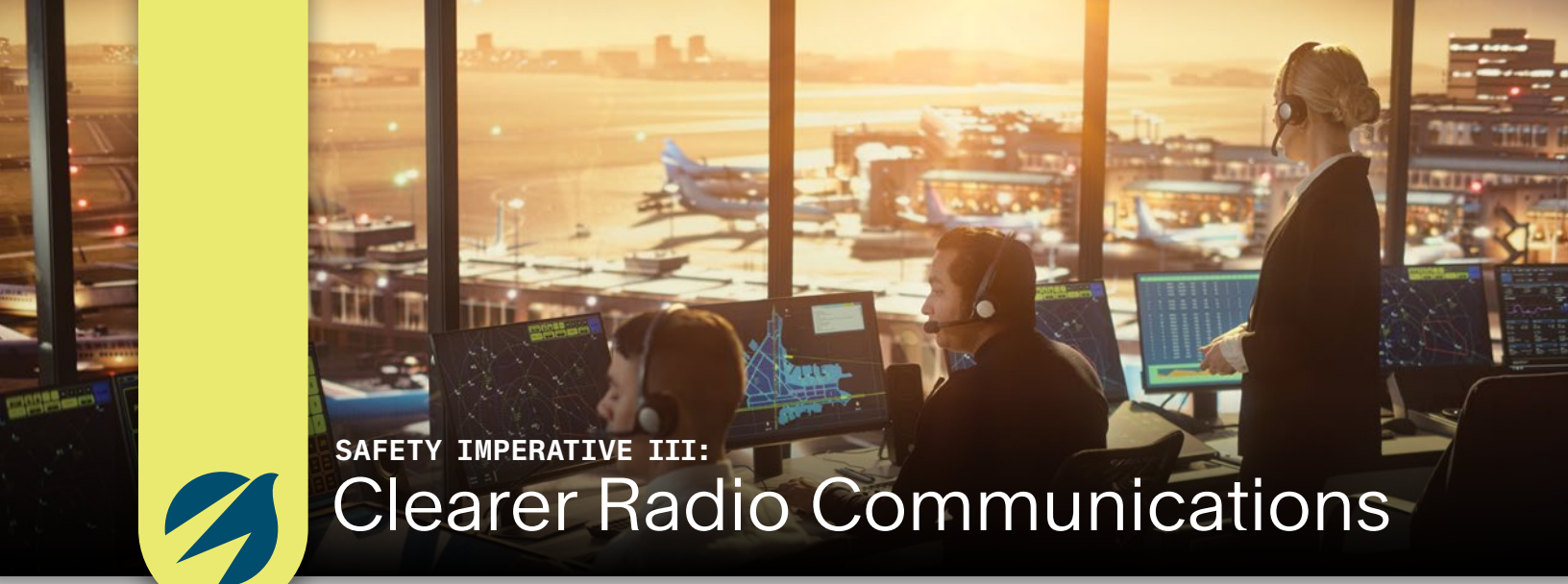
[19]Seoane, Steve. (2019). *How wearable tech can help monitor first responders' health*. EMS1. https://www.ems1.com/ems-products/computer-aided-dispatch-cad/articles/how-wearable-tech-can-help-monitor-first-responders-health-imJkMAIYuWvs9loF/

[20]Niotis K, Saif N, Simonetto M, Wu X, Yan P, Lakis JP, Ariza IE, Buckholz AP, Sharma N, Fink ME, Isaacson RS. *Feasibility of a wearable biosensor device to characterize exercise and sleep in neurology residents*. Expert Rev Med Devices. 2021 Nov;18(11):1123-1131. https://pubmed.ncbi.nlm.nih.gov/34632903/

[21]Ghosh K, Nanda S, Hurt RT, Schroeder DR, West CP, Fischer KM, Bauer BA, Fokken SC, Ganesh R, Hanson JL, Lindeen SA, Pruthi S, Croghan IT. *Mindfulness Using a Wearable Brain Sensing Device for Health Care Professionals During a Pandemic: A Pilot Program.* J Prim Care Community Health. 2023 Jan-Dec;14. https://pmc.ncbi.nlm.nih.gov/articles/PMC10041582/

MERLIN

HEADQUARTERS: 129 South Street, Boston, MA 02111



SAFETY IMPERATIVE III:

Clearer Radio Communications

Effective communication between pilots and air traffic controllers is critical for maintaining safety in the NAS. However, miscommunications—such as missed and misunderstood radio calls—are regularly highlighted as an ongoing challenge in commercial aviation[22].

Between January and October 2023, the Federal Aviation Administration (FAA) identified 19 serious runway incursion incidents, the highest in any 10 month period since 2016[23]. These incidents, many of which involved radio miscommunications, were an impetus for the FAA to charter an independent expert National Airspace System Safety Review Team to research factors contributing to runway incursions and make specific recommendations. And in 2024, continued near-miss incidents caused the FAA to open an audit into runway incursion risks at 45 of the United States' busiest airports[24].

While the cause of these runway incursions is complex and multifaceted, one key opportunity highlighted in the final report issued by the National Airspace System Safety Review Team[25] is more effectively leveraging infrastructure and technology to meet the needs of the modern NAS. Additionally, the National Transportation Safety Board (NTSB) stated in November 2023 that "while [serious near-miss] events are incredibly rare, our safety system is showing clear signs of strain that we cannot ignore"[26]. Both of these examples underscore that now is the time to re-imagine and invest in technologies that can help prevent future incidents. More specifically, automation that assists with communications can play a key role.

Merlin is developing a communication assistance system that leverages the modern benefits of NLP, a technology that can translate radio transmissions in real-time and provide context-aware responses. Such systems listen to radio communications and make recommendations to pilots if a transmission is relevant for their aircraft. Additionally, these systems process the transmissions of other aircraft and can understand the position and intent of those vehicles' operations. Using this information, it is possible to build a comprehensive situational awareness picture of an aircraft's operating environment, which can facilitate the pilot's formation of their own mental model of the scenario. Merlin envisions such systems can operate in an assistive capacity, similar to how

[22]Federal Aviation Administration (FAA). (2024). *Runway safety statistics*. U.S. Department of Transportation. https://www.faa.gov/airports/runway_safety/statistics
[23]SimpleFlying. (2023). *FAA Identifies 19 Serious Near Miss Incidents In 2023: The Most In 7 Years*. https://simpleflying.com/faa-near-miss-incidents-2023/
[24]FAA. (2024). *General statements*. U.S. Department of Transportation. https://www.faa.gov/newsroom/statements/general-statements
[25]FAA. (2023). *NAS safety review team report*. U.S. Department of Transportation. https://www.faa.gov/NAS_safety_review_team_report.pdf
[26]Shepardson, David. (2023). *NTSB chair says US near-miss aviation incidents 'clear warning sign'*. Reuters. https://www.reuters.com/business/aerospace-defense/ntsb-chair-says-us-near-miss-aviation-incidents-clear-warning-sign-2023-11-08/.

MERLIN

HEADQUARTERS: 129 South Street, Boston, MA 02111

Ground Proximity Warning Systems (GPWS) and Traffic Alert and Collision Avoidance Systems (TCAS) are already used by pilots and have delivered dramatic safety benefits.

Developing technology that can understand an aircraft's operational environment is fundamental to advancing automation in aviation, as this awareness allows systems to alert pilots when the state of their aircraft may be at risk. For example, an NLP system could trigger an alert if it observes an aircraft crossing the runway hold short line at an airport, but does not hear a transmission from air traffic control clearing that aircraft to actually enter the runway environment. Similarly, the system could warn pilots if they're trying to land on a runway without actually having received a landing clearance.

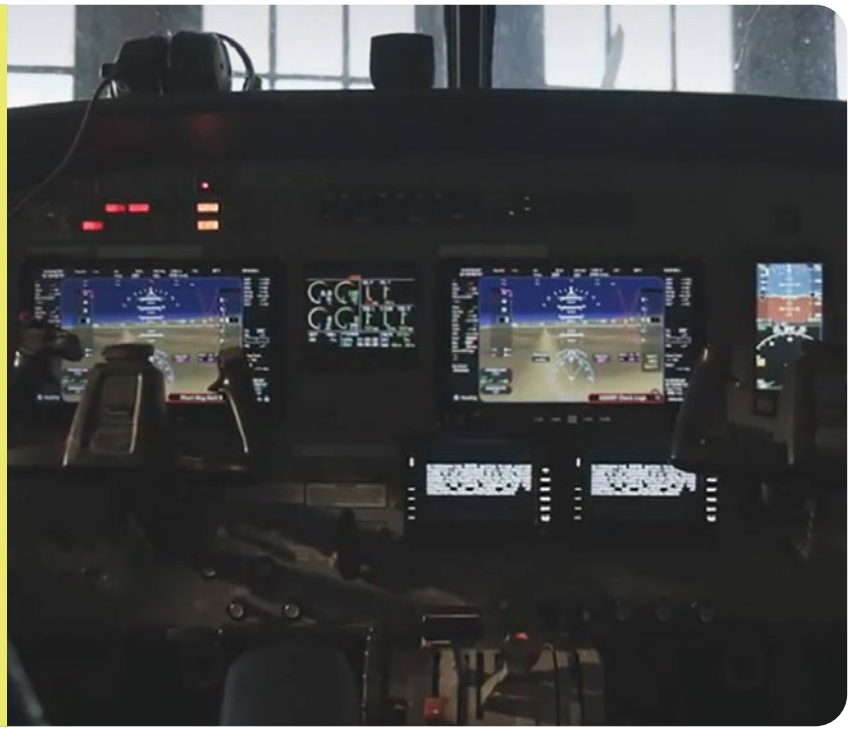
NLP and similar technologies could also be extended to make existing systems more robust. For example, an "enhanced" traffic collision avoidance system could leverage what is actively being broadcast over the radio in conjunction with its existing sensors and capabilities. Imagine a traffic pattern scenario at a non-towered airport where non-cooperative traffic is more common. If a pilot provides a position report over the radio, but their aircraft is not visible via Automatic Dependent Surveillance–Broadcast (ADS-B) or other cooperative methods, the system could focus its cameras or other visual sensors in the area of the pilot's position report in an attempt to locate the traffic.

Such technology also has applications beyond the flight deck. For example, a system that actively monitors radio calls and compares those calls to actual aircraft positions would be just as useful to an air traffic controller as it would be to a pilot, and would likely deliver a similar level of safety benefit if installed in an air traffic control facility.

Automated Systems to Reduce Radio Miscommunications

By viewing radio communications as a skill and developing systems that enable operators to better perform that skill, Merlin is demonstrating how the existing strengths of human pilots can be paired with automation to unlock new levels of safety and performance.

Our communication assistance system leverages the modern benefits of NLP, a technology that can translate audio in real-time and provide context-aware responses, to increase pilot situational awareness and reduce missed or misunderstood radio transmissions.



MERLIN

HEADQUARTERS: 129 South Street, Boston, MA 02111

Recommendations

Advancing aviation safety requires motivated cooperation between government and industry. The FAA, in particular, has a unique opportunity to promote advancements in automation and help usher in a new era of innovation for the industry. We recommend the following actions to aid the development of automation technologies in ways that will drive increased aviation safety:

RECOMMENDATION I

Create Aviation Communication Datasets for NLP Training

Developing NLP algorithms requires large datasets of existing aviation communications that can be used to train and verify models. These datasets must capture the breadth and variety of today's aviation communications, and creating them is inherently a collaborative effort between regulators, industry, and academia to ensure that results are comprehensive, sufficient, and impartial.

The FAA should launch an initiative to create publicly available datasets of aviation communications for the purposes of advanced technology development. Such an effort would be similar to how the FAA previously worked collaboratively with MIT Lincoln Laboratory and broader industry to develop airspace encounter models, which subsequently supported standards creation and development of collision avoidance systems[27].

Specific Actions Required

1. Identify a research institution to lead dataset development.
2. Convene an industry working group to provide stakeholder input during development.
3. Make large amounts of aviation communications data available to relevant stakeholders to support dataset creation.
4. Create and release the datasets according to a to-be-determined timeline.

[27]https://github.com/Airspace-Encounter-Models

MERLIN

HEADQUARTERS: 129 South Street, Boston, MA 02111

Key Involvement

- FAA: Charter the effort and organize stakeholders; define broad schedule and timelines.
- Research Institution: Lead dataset development and release.
- Industry Working Group: Provide customer input during dataset planning and development, and offer feedback once initial dataset versions are available.

Success Metric

- Plan and execute a trial project that demonstrates the datasets successfully interpreting a broad variety of aviation communications, while meeting pre-defined performance metrics.

Anticipated Challenges

- Data Availability: Records of air traffic control and/or other aviation communications are not widely available; the FAA and project stakeholders will need to determine methods by which large amounts of such data can be sourced and used for dataset creation.
- Data Privacy: Source data should be appropriately anonymized to ensure stakeholder trust.

Establishing an effort to create training datasets for aviation communications would immediately accelerate the development of systems that enhance safety within the NAS, as well as inform standards and other policymaking efforts to formalize verification and validation of NLP algorithms going forward.

RECOMMENDATION II

Quantify Aviation Communication Performance Metrics

Developing systems that can enhance aviation communications, and ultimately verifying their effectiveness, first requires an understanding of the performance of existing aviation communications. However, there have been few recent, large-scale efforts to quantify performance metrics and error rates associated with very high frequency (VHF) communications in the modern NAS.

The FAA should charter an effort to understand the performance of existing VHF communications. By quantifying metrics such as how often flight crews and air traffic controllers issue corrections or how frequently radio transmissions are blocked, the industry can identify inefficiencies and improve communication protocols.

Specific Actions Required

1. Identify a research institution to lead the effort.
2. Convene an industry working group to define performance metrics and methods by which these metrics should be measured.
3. Make large amounts of aviation communications data available to relevant stakeholders.
4. Draft and release a comprehensive report providing aviation communication performance metrics, in consultation with stakeholders as required.

Key Involvement

- FAA: Charter the effort and organize stakeholders; define broad schedule and timelines.
- Research Institution: Lead research effort and final release of comprehensive report.
- Industry Working Group: Provide stakeholder feedback and input as part of the research effort.

Success Metrics

- Release a comprehensive report detailing VHF communications performance metrics according to a to-be-determined timeline.

Anticipated Challenges

- Data Availability: Records of air traffic control and/or other aviation communications are not widely available; the FAA and project stakeholders will need to determine methods by which large amounts of such data can be sourced and used for performance analysis.
- Data Privacy: Source data should be appropriately anonymized to ensure stakeholder trust.

Baselining the performance of existing aviation communications is essential to understanding how and where the system can be improved. By chartering a report detailing such information, the FAA can provide a key input that will ultimately lead to improving safety in the NAS.

RECOMMENDATION III

Streamline a Pathway for Integrating AI Technologies

Currently, the FAA requires any applicant using AI on an aircraft to undergo the Issue Paper process, regardless of the safety criticality of the AI[28]. Issue Papers are a well-established and useful concept in aviation certification, particularly when a well-established means of compliance does not exist. However, the position that any use of AI immediately triggers the Issue Paper process is overly burdensome in the context of AI used in functions with only low safety criticalities.

By definition, non-safety critical applications are not considered a risk for contributing to failure conditions, and thus do not typically require additional above-and-beyond measures for proving safety and compliance. The net effect of the FAA's requirement is that undue burden is placed on using AI in low criticality applications, which further dissuades applicants from employing the use of AI altogether.

If conventional compliance methodologies were followed, this burden wouldn't exist. In fact, the regulator would not normally be aware AI was even being used, because applications with only low safety criticalities do not require complete traceability below their highest level requirements. As such, their means of implementation would not be known.

This Issue Paper requirement is especially problematic in the context of Non-Required Safety Enhancing Equipment (NORSEE), a construct the FAA has promoted to integrate non-required safety enhancing technologies into aircraft flight decks with streamlined certification, provided they operate on a non-interference basis[29]. NORSEE provides the ideal framework for introducing new technologies in non-safety critical roles, and has already proven successful at allowing for safety-increasing technologies to become operational faster. A prime example is the integration of electronic flight bags (EFBs) into aircraft cockpits, which are now recognized as a key asset in improving pilot situational awareness. However, because NORSEE technologies using AI are required to go through the Issue Paper process, a framework that is normally meant to spur innovation cannot be leveraged as intended.

Going forward, the FAA should adjust its guidance to no longer require additional compliance showing activities, including Issue Papers, for AI used in non-safety critical applications. Further, NORSEE should be encouraged for applications where a non-required safety benefit can be delivered using AI, not only in the areas specified in NORSEE's guidance document[30], but including operations under 14 CFR Part 25. Adjusting this guidance would streamline the process for introducing novel technologies in non-safety critical roles, which then provides a platform for testing and refining innovations in a real operational environment. Consequently, novel technologies could be introduced and operationalized faster, while maintaining the existing high standards of safety in aviation.

[28]Federal Aviation Administration (FAA). (2024). *Small Airplane Issues List*. Item A-1601. U.S. Department of Transportation. https://www.faa.gov/sites/faa.gov/files/2024-09/small_airplane_issues_list_2024_q3.pdf,

[29]Federal Aviation Administration (FAA). (2016). *Policy Statement PS-AIR-21.8-1602*. U.S. Department of Transportation. https://drs.faa.gov/browse/excelExternalWindow/1790B02F1833357486257F9200592110.0001

[30]Federal Aviation Administration (FAA). (2016). *Policy Statement PS-AIR-21.8-1602*. U.S. Department of Transportation. https://drs.faa.gov/browse/excelExternalWindow/1790B02F1833357486257F9200592110.0001

Specific Actions Required

1. Adjust FAA guidance to relieve the current requirement for additional, burdensome compliance activities for AI used in non-safety critical applications.
2. Charter a standards organization or industry working group to draft additional guidance and/or language that would clarify the use of AI within the NORSEE framework.
3. Circulate the draft language with industry stakeholders, academia, and regulators for feedback as required via established processes.

Key Involvement

- FAA: Adjust its internal guidance as discussed.
- Standards Organization: Lead effort to publish additional guidance for AI implemented via the NORSEE framework.
- Industry Stakeholders: Provide feedback and input as needed.

Success Metrics

- Release of updated language and/or guidance regarding the use of AI, machine learning, and similar development methodologies implemented via NORSEE.

Anticipated Challenges

- FAA Resource Limitations: Personnel and funding limitations at the Agency can be mitigated by leveraging external standards bodies or other trusted stakeholders to provide initial draft language and suggestions for improved NORSEE guidance, in consultation with industry stakeholders.

Issuing clear guidance detailing how systems that leverage AI and machine learning can be included within the NORSEE framework, without the need for Issue Papers or other burdensome processes, would provide a method for additional learnings and demonstrations of how these systems can further increase aviation safety.

Conclusion

The next chapter in commercial aviation safety won't be written by replacing human expertise with automation—it will come from finding better ways for pilots and technology to work together. As we've outlined, there are immediate opportunities to enhance safety through smarter workload distribution, better fatigue management, and clearer radio communications. These are incremental improvements, and they're low-lift examples of how we can create an even safer flying experience for crew and passengers alike.

The path forward requires a balanced approach. We must leverage new technologies like natural language processing and biometric monitoring while staying true to commercial aviation's proven safety track record. This means working closely with pilots, regulators, and industry partners to ensure new automaton tools enhance rather than compromise human judgment and decision making.

At Merlin, we see these challenges as opportunities. By focusing on human-centered automation that augments pilot capabilities, we can address critical safety imperatives while laying the groundwork for future advances. Our work with military and commercial partners demonstrates that meaningful progress is possible when we take a thoughtful, incremental approach to innovation. Through our recommended framework—creating robust training datasets, establishing clear performance metrics, and streamlining the integration of AI technologies—we can accelerate the development of safety-enhancing automation while maintaining aviation's high standards. By embracing this vision and working together across industry and government, we can ensure that the next generation of flight operations is the safest yet.

CONNECT:

MEDIA INQUIRIES:

Request Media Kit: media@merlinlabs.com

MERLIN

HEADQUARTERS: 129 South Street, Boston, MA 02111