FLIGHT SAFETY DATA STORYTELLING: CONTINUOUS LEARNING FROM WHAT WENT RIGHT

Barth Baron, Jr.
San José State University
San José, California

Melissa Peterson
San José State University
San José, California

The anticipated increase in air traffic, airspace complexity, and safety data volume inspired the development of an In-time Aviation Safety Management System (IASMS). Evolving from the legacy Safety Management System (SMS), IASMS will use machine learning (ML) to improve current risk management and safety assurance methodologies, predict future threats, and reduce analysis and response times. We propose applying the IASMS concept to directly support skilled performance in real-time with an interactive safety visualization tool for pilots. This tool would analyze disparate data sets to reveal latent operational threats and provide examples of skilled performance adaptations. Pilots would access tailored ML-enabled predictive data narratives through a user-centered interface. Challenges with using ML in this context, data integration, and interface design are discussed.

This paper proposes accelerating the expertise of flight crews by giving them access to interactive graphical data narratives from previous crews in analogous situations. The goal is to have crews draw on ML-generated predictions in real-time to anticipate operational threats based on current and forecast conditions, monitor for their possible occurrence, and be informed by adaptive behaviors of previous crews. Such proactive behaviors are often latent in the safety data, hidden in flights without incident reports, which this system would access and display using ML algorithms. This idea requires combining lessons from user experience (UX) design, cognitive systems engineering (CSE), and data science to produce real-time safety data narratives for pilots, capturing where normative models require pilots to exhibit adaptive behaviors. We propose using ML on diverse data sources, similar to the In-time Aviation Safety Management System (IASMS), where advanced analytics will synthesize large and diverse datasets to create predictive safety capabilities (Ellis et al., 2022). In some cases, a pilot will pull from the system to update their understanding of the situation. In other cases, the system will push updates about evolving threats and supply possible mitigations. The practical and theoretical underpinnings of this system and its design are discussed here.

Supporting Resilient Performance

While previous successes in improving safety were accomplished by analyzing accidents and incidents to prevent their recurrence, some recent research focuses on analyzing successful outcomes to understand resilient flight crew performance (Hollnagel et al., 2015). The Resilience Engineering (RE) framework developed by Hollnagel et al., with four phases of performance summarized as “anticipate, monitor, respond, learn,” applies to both the system and individual levels. As currently described, IASMS focuses on predictive safety at the system level, leaving a gap for IASMS capacities to support resilient performance on the individual level. Previous explorations of resilient safety behaviors have included human-interpreted case studies and Aviation Safety Reporting System (ASRS) and Aviation Safety Action Program (ASAP) reports. However, this process requires absorbing substantial time and attentional resources from managers. It also relies on the pilot reports having supplied sufficient narrative detail to help the analyst fully extract the productive safety behaviors. For these reasons, newer research has begun to use ML to explore these datasets.

Machine Learning to Discover Latent Safety Data

Anticipated changes in the National Airspace System (NAS) inspired the proposal of a next-generation SMS based on predictive analuytics (Ellis et al., 2022). Reflecting the need to manage safety by predicting threats in real-time and on a system level, the proposed IASMS will leverage ML to create a proactive safety framework (National Academies of Sciences, Engineering, and Medicine et al., 2018).

To support developing the predictive capacities needed for IASMS, a novel approach is currently exploring the use of ML to analyze ASRS reports for proactive safety behaviors in data that might not otherwise reveal resilient performance. (Matthews et al., 2023). This also has the benefit of revealing the conditions under which such adaptations are required. Using a natural language processing (NLP) algorithm to extract evidence of positive flight crew behaviors in narratives that described events with negative outcomes, Matthews et al. demonstrated the concept of repurposing incident data to generate representations of adaptive performance.

One component the proposed system is applying a similar approach to Flight Operations Quality Assurance (FOQA) data, training an ML algorithm to find patterns of operational and environmental conditions associated with unintended aircraft states, and learning the adaptive performance of crews that experience analogous conditions without incident. This starts by generating a comprehensive characterization of a flight's environmental and operational variables where an unintended aircraft state occurs, flagged by a FOQA exceedance, ASAP, or ASRS report. The algorithm is then trained on situations with equivalent conditions where no exceedance occurred to learn what adaptations are associated with non-event flights.

Data Storytelling

The second component of this proposal is making these characterizations of adaptive performance available for pilots in real-time as they approach analogous operational and environmental conditions. This allows pilots to anticipate threats and apply adaptive behaviors learned from the algorithm. The goal is to enable crews to monitor for and respond efficiently to threats the system anticipates. Bringing these algorithm outputs to pilots in a meaningful form is a significant challenge, to which we now turn our attention.

Data tables, visualizations, and, most recently, dynamic data dashboards have been used to communicate insights from data. However, there is some evidence that such visualizations and dashboards can lead to high cognitive load without high levels of visualization literacy (VL). Since adding more cognitive load in the flight deck is counterproductive to managing abnormal events, the system imagines redesigning the visualization into data-storytelling visualizations (Liu et al., 2024). This method provides a way to package complex data into an easily digestible format for immediate use (Oberascher et al., 2023). This technique also puts the viewer into the role of a doer – data storytelling can provide important context, insights, and a call to action (Lo Duca & McDowell, 2024). The need for narrative-based insight is further supported by the work done on ML-supported software tool Kaona, which curates narratives for users from ASRS data (Paradis et al., 2025). The current project differs from Kaona in that these narratives would be further simplified for pilots to make quick decisions and supplemented with interactive visualizations.

System Design

We propose designing a platform for front-line operators that connects with a ground-based system responsible for integrating the diverse safety-related datasets, including real-time traffic and meteorological data. The airborne component adds its atmospheric data and current aircraft state fused from aircraft sensors. As aircraft connectivity and computing power matures, onboard systems will be able to access outside data sources to form its own understanding of how the situation is evolving, allowing integration into the analysis. The proposed interface converts data to contextually driven data narratives and generates recommendations based on successful outcomes from other flights experiencing similar situational variables. A use case and potential interface designs are described here.

Simple Example Use Case

During every arrival, pilots face a potential threat related to speed limits for the wing flaps. For example, Airbus aircraft have four flap settings: Flaps 1, Flaps 2, Flaps 3, and Flaps Full. Each flap setting has a maximum airspeed, exceeding which triggers a warning and can lead to aircraft structural damage. Each setting also has a minimum safe speed to provide a margin above aerodynamic stall, and the next setting must be selected before decelerating below that minimum speed. Therefore, there is a window of speed operation between the upper and lower limits. Landing gear can be extended to create drag, aiding deceleration, and the landing gear also has a maximum extension speed. Training emphasizes a configuration sequence of selecting Flaps 1, decelerating until the minimum Flaps 2 speed, then extending the landing gear and selecting Flaps 3 when further deceleration is required, before selecting Flaps Full for final deceleration and landing. This is the normative model taught in training.

Consider the scenario of an Airbus operating at a heavy weight that is cleared for an approach, intercepting the glide slope at 5,000 ft and 190 kts at Flaps 1 with a tailwind. The tailwind causes a higher-than-normal ground speed, which in turn causes a higher-than-normal rate of descent at glideslope intercept, causing the airspeed to increase. The crew, adhering to a normative work model, configures to Flaps 2. However, the environmental conditions cause the airspeed to increase beyond the Flaps 2 limit speed, triggering a FOQA event and requiring an aircraft inspection. The crew may also voluntarily submit a confidential Aviation Safety Action Program (ASAP) incident narrative. However, another crew in the same conditions configures out of sequence, extending the landing gear with Flaps 1. The extra drag overcomes the acceleration as the flight descends on the glideslope, allowing deceleration to a point where Flaps 2 can be safely selected. This example of descriptive work avoids exceeding aircraft design constraints, however the adaptation goes undetected by safety intelligence systems.

Learning From What Went Right in Other Flights: Turning FOQA Data Into Examples of Skilled Performance

In an example data-fusion solution, the FOQA event of a flap overspeed triggers the ML algorithm to generate a representation of the incident aircraft's state during the event. Having established the environmental, control, and situational variables, the algorithm searches the FOQA database for cases where flights started with analogous initial conditions but did not trigger a FOQA event. The system then queries diverse datasets, ranging from atmospheric stability, proximity to other aircraft, learning what work prevented the undesirable outcomes in the specific operational context. In the above example, the crew diverged from the descriptive model by extending the landing gear out of the sequence taught in training. The system can learn multiple pathways to attaining higher-order functional goals outside the normative model. The challenge remains of informing pilots of the operational threat's potential and providing access to these adaptive pathways without increasing cognitive load.

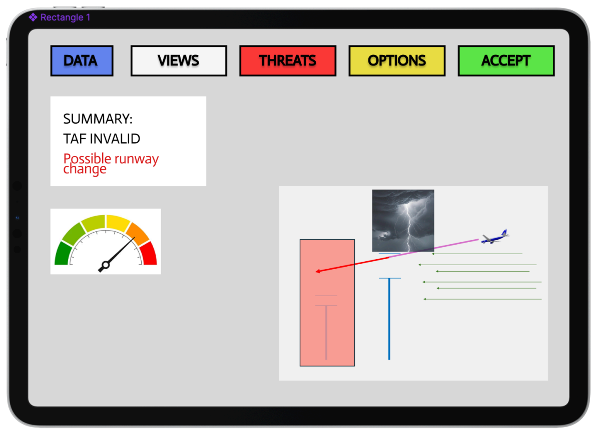
Interface Design

We propose connecting pilots with the outputs of the system's ability to fuse and analyze data using ecological interface design theory (EID) (Flach & Bennett, 2017; Vicente & Rasmussen, 1992; Vicente, 2002). This interface would be developed using iterative design process (Savage, 1996) to build an interface that allows crews to receive notifications of emerging threats generated by the system. Additionally, pilots can inform the system of the flight's current operational goal and query the integrated dataset. In both cases, the crew receives an interactive visualization of potential threats and examples of successful mitigations tailored to the goals and constraints of the operation at hand. EID is particularly appropriate as it integrates the complex system hierarchy of the airplane and surrounding variables for the end user, to support near real-time expertise transfer and formative work development. This visualization should include markers of where success may require going beyond standard procedures and representations of the potential results of not adapting the descriptive model. At this early stage of development, an iterative design process is necessary to meet design best practices, ensure usability, and limit cognitive load. The descriptions and images provided here constitute an initial starting point for this process.

A tablet interface, connected through high-speed internet to the ground-based ML data fusion and analysis system, continuously looks at current conditions, telling a story about emerging threats and offers potential mitigations (see Figure 1). The system will hold initial flight conditions, such as aircraft type, weight, location, time, destination, and Flight Management Computer (FMC) data. Environmental factors such as weather and other traffic are automatically updated to frame the situation comprehensively and continuously. The output is tailored to address the current operational constraints and provides solutions to problems experienced in the conditions they're currently facing.

Figure 1

Mockup of an interactive situation interface.



Note. 3-D integrated image of the situation and emerging threats to be interactive and manipulated by the user.

Challenges

This project is in the early conceptual stages and will rely on technological advances as it progresses. It requires integrating disparate data streams comprising narrative and numerical data, which have not already been normalized for combination. This is an important step, as the data sources (ASRS, ASAP, FOQA) contain potential solutions for a crew confronting threats.

Beyond data acquisition and normalization, a further challenge is using machine learning and big data in the flight deck environment, which is currently not connected to reliable high-speed internet or servers that drive data integration and processing. Given the need for real-time processing of the current flight variables, predicting future threats, and recommended mitigation strategies, the traditional methods of batch processing and algorithm iteration are inappropriate. Instead, stream-based and online machine learning techniques are required (Andreoni Lopez et al., 2019; Medeiros et al., 2020), or, in the shorter term, providing the onboard algorithm with a flight-specific subset of data to be used statically until landing. This problem will become less impactful as in-air high-speed connections become more common, allowing flights to access the data and ML outputs more quickly.

While much of the technology discussed here is still in the early or conceptual stages, specific intermediate steps, such as testing subsets of data integration and prototyping displays, could allow for early testing of the concepts outlined here. This will also allow for proof-of-concept exploration in simulators, as the final interface can be iterated using models from current IASMS efforts before requiring live updates and modelling over high-speed data connections. We submit that these steps should be considered concurrently within the larger IASMS framework, making pilots key stakeholders in implementing safety intelligence.

References

Andreoni Lopez, M., Mattos, D. M. F., Duarte, O. C. M. B., & Pujolle, G. (2019). Toward a monitoring and threat detection system based on stream processing as a virtual network function for big data. *Concurrency and Computation: Practice & Experience*, *31*(20), e5344.

Ellis, K. K., Prinzel, L. J., Krois, P., Davies, M. D., Oza, N., Stephens, C., Mah, R., Infeld, S. I., & Koelling, J. H. (2022, June 27). A future in-time aviation safety management system (IASMS) perspective for commercial air carriers. *AIAA AVIATION 2022 Forum*. AIAA AVIATION 2022 Forum, Chicago, IL & Virtual. https://doi.org/10.2514/6.2022-3220

Flach, J. M., & Bennett, K. B. (2017). Improving sensemaking through the design of representations. In *Cognitive Systems Engineering* (pp. 165–180). CRC Press.

Hollnagel, E., Wears, R. L., & Braithwaite, J. (2015). *From safety-I to safety-II: A white paper*.

Liu, Y., Pozdniakov, S., & Martinez-Maldonado, R. (2024). The effects of visualisation literacy and data storytelling dashboards on teachers' cognitive load. *Australasian Journal of Educational Technology*. https://doi.org/10.14742/ajet.8988

Lo Duca, A., & McDowell, K. (2024). Using the S-DIKW framework to transform data visualization into data storytelling. *Journal of the Association for Information Science and Technology*. https://doi.org/10.1002/asi.24973

Matthews, B., Barshi, I., & Feldman, J. (2023). An approach to identifying aspects of positive pilot behavior within the aviation safety reporting system. *2023 IEEE/AIAA 42nd Digital Avionics Systems Conference (DASC)*, 1–6.

Medeiros, D. S. V., Cunha Neto, H. N., Lopez, M. A., S. Magalhães, L. C., Fernandes, N. C., Vieira, A. B., Silva, E. F., & F. Mattos, D. M. (2020). A survey on data analysis on large-Scale wireless networks: online stream processing, trends, and challenges. *Journal of Internet Services and Applications*, *11*(1). https://doi.org/10.1186/s13174-020-00127-2

National Academies of Sciences, Engineering, and Medicine, Division on Engineering and Physical Sciences, Aeronautics and Space Engineering Board, & Aviation Safety Assurance Committee. (2018). *In-time aviation safety management: Challenges and research for an evolving aviation system*. National Academies Press.

Oberascher, L., Ploder, C., Spiess, J., Bernsteiner, R., & van Kooten, W. S. M. T. (2023). Data storytelling to communicate big data internally – a guide for practical usage. *European Journal of Management Issues*. https://doi.org/10.15421/192303

Paradis, C. V., Hong, C., Matthews, B., Davies, M. D., & Hooey, B. (2025). Kaona: Deep searching and curating aviation safety reporting systems. *AIAA SCITECH 2025 Forum*. https://doi.org/10.2514/6.2025-2152

Savage, P. (1996). User interface evaluation in an iterative design process: a comparison of three techniques. *Conference Companion on Human Factors in Computing Systems Common Ground - CHI '96*. Conference companion, Vancouver, British Columbia, Canada. https://doi.org/10.1145/257089.257335

Vicente, K. J., & Rasmussen, J. (1992). Ecological interface design: theoretical foundations. *IEEE Transactions on Systems, Man, and Cybernetics*, *22*(4), 589–606.

Vicente, Kim J. (2002). Ecological interface design: progress and challenges. *Human Factors*, *44*(1), 62–78.