

Open-Source Flight Data: Free, but at What Cost? [†]

James Norman 

John D. Odegard School of Aerospace Sciences, The University of North Dakota, Grand Forks, ND 58202, USA; james.e.norman@und.edu

[†] Presented at the 10th OpenSky Symposium, Delft, Netherlands, 10–11 November 2022.

Abstract: The proliferation of open-source flight data has been heralded throughout the aviation industry, from air traffic service providers seeking efficiencies to the voyeuristic tracking of celebrity aircraft movements. However, the rise of open-source data presents challenges when juxtaposed against the traditional safety management systems that have sustained commercial aviation until now. Open-source data are not protected legally to the same extent as traditional flight data, and data from open-source platforms are not subject to the same validation standards as traditional flight data. Pilots are now able to interface with their own flight data, which diverges from the original aggregate intent of flight data monitoring programs. This paper addresses regulatory, legal, and ethical topics pertaining to the use of open-source data in U.S. commercial aviation, and offers recommendations to the industry in light of the new data landscape.

Keywords: ADS-B; data ethics; flight data monitoring (FDM); flight operations quality assurance (FOQA); safety management systems; voluntary safety programs

1. Introduction

Issues surrounding data ownership and use in modern-day society are becoming increasingly important to society in general, and to aviation in particular. Be it privacy, security, personal agency, or a myriad of other issues, those participating in the electronically interconnected world are finding that the rate of technological development outpaces the human capacity to comprehend complex and opaque systems. Technology theorists have proposed a societal shift to netizens [1], where even basic human-to-human interaction is predicated on the use of technology: computers, mobile phones, social networking, etc. Therefore, data have become arguably the most essential element of contemporary human existence.

The commercial aviation domain provides an excellent case study for this issue. Until quite recently, the flow of flight data was tractable, predictable, and well-understood by stakeholders. After an aircraft completed a flight, the relevant data were uploaded to a server, validated, and analyzed, and the outcomes were shared with the pilots, company, and regulator as part of the organization's safety management system. This formalized process was memorialized in contracts between labor and companies, and built on guidance from the regulator.

Today however, a new data source has been introduced into the aviation ecosystem: open-source data. This type of data may take on many forms, including air traffic voice data, weather, and trajectory-based data emanating from the aircraft. The ease of obtaining open-source data has had a democratizing effect; knowledge that was once proprietary is no longer. For example, freedom of information requests to obtain voice recordings are often no longer necessary, as they are available on a popular aviation website. Similarly, flight data are no longer shielded by the airlines, as many parameters are available from ADS-B.

The flattening of the aviation data playing field has brought about new opportunities for academic research, air traffic management, sustainability efforts, and overall aviation



Citation: Norman, J. Open-Source Flight Data: Free, but at What Cost? *Eng. Proc.* **2022**, *28*, 9. <https://doi.org/10.3390/engproc2022028009>

Academic Editors: Michael Schultz and Junzi Sun

Published: 16 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

safety improvement. Novel use cases such as the tracking of celebrity aircraft and the whereabouts of CEOs can offer insights that have potential ramifications beyond aviation, and can potentially affect financial markets [2]. Estimation of aircraft emissions from open-source data may help regulators and non-government organizations better understand climate issues without having to rely on data disclosure from contributing entities.

Yet, as with most developing technology use cases, there may be unintended consequences. One area in which open-source data may have unintended consequences relates to the voluntary safety programs that have developed in the United States during the past two decades. These formalized and consistent programs provide vital safety management data to airlines. However, the assumptions and protections these programs are built upon are being challenged by the rise of open-source data.

The goal of this paper is to raise awareness of issues pertaining to both open-source data and proprietary data used in novel settings as they relate to voluntary safety programs administered by the FAA in the United States. Left unaddressed, these issues may develop into challenges that could put well-intended safety programs in jeopardy. A worse-case scenario may be a regression of the quality, just culture, and trust that have developed within U.S. commercial aviation. Three issues form the basis of this inquiry: the partnership paradigm, the scope and protection of safety data, and the ethics of big data use in aviation.

While not exhaustive, these important topics offer a triangulated approach that may add value for the industry and regulators as aviation data takes on increasingly varied dimensions. The scope of this paper pertains exclusively to the U.S. Therefore, it is acknowledged that the aspects of data use and protection discussed herein may not be applicable or transferable to other states. However, the principles of just culture and ethical use of data should transcend national boundaries in light of aviation's global presence and efforts for standardization.

2. Background

To understand the interplay between open-source data and traditional flight data monitoring programs in the U.S., a brief overview is necessary. Flight operational quality assurance (FOQA) started nearly 20 years ago as part of a quartet of voluntary safety programs codified by the FAA. While in theory voluntary, FOQA is present at all U.S. airlines due to requirements of an SMS, insurers, and ICAO standards and recommended practices (SARPs). The FAA's guidance document for FOQA programs has remained unchanged from its creation. Its tenets include the aggregate use of flight data and the goal of spotting and mitigating adverse trends before they become incidents or accidents.

In FOQA programs, certain persons from the employee group are designated as gatekeepers. The gatekeeper is "the individual(s) who can link FOQA data to an individual flight or crewmember" [3]. The role of the gatekeeper has been a key to the trust, success, and proliferation of FOQA programs. The pilots working in these programs are not emissaries of the company; rather, they are chosen by the pilots' union or employee group. They hold no disciplinary power. The pilot gatekeeper is empowered to make decisions on the use, scope, and extent of flight data use.

The gatekeeper has historically had a unique role in flight safety, namely, the ability to make personal contact with a pilot regarding their data when necessary. Seen as a peer, the gatekeeper is able to obtain supplemental information because of the trusted and protected nature of the conversation. This is a win for all parties: the pilot feels comfortable and protected sharing sensitive information, the labor or employee representative has a better window into the operation, and the company receives richer and more meaningful information, thereby enhancing safety.

The intention of FOQA for aggregate data analysis and the use of the gatekeeper as a firewall between identified flight data and the company has been disrupted by tablet devices and the newfound ability to provide personalized flight data information to pilots. A multitude of companies have entered into agreements with airlines to provide this type of

data. Pilots can now review animations of their landing, isolate data such as the touchdown point, and track their performance against benchmarks among their fleet or airline.

The use of a personal tablet to obtain flight data along with its non-aggregate nature potentially bypasses the role of the gatekeeper and perceptually shifts the intention of a FOQA program from the aggregate to the singular. In this new world, a gatekeeper has less ability to add context and nuance when analyzing data or making a crew contact; the pilot can draw conclusions on their own based on their flight data app. The company has the potential to bypass the expertise of the gatekeepers and labor group, and create measurements and definitions of exceedance events on their own terms in a way that is devoid of the collaborative nature that has brought the industry to where it is today.

3. The Partnership Paradigm

The FAA terms its voluntary safety programs “partnership programs” [4]. This intentionally sets a tone of collaboration and common cause between the company, FAA, labor groups, and other industry stakeholders. Partnership connotes equality and transparency; however, this parity may be challenged when one party is privy to more information than the other. With the ease of obtaining open-source data, stakeholders may develop differing views of the landscape. Because open-source data are not processed through the traditional safety management system, they can potentially lack the validation that flight data management programs rely upon.

The issue of how to treat proprietary versus open-source data is a potential source of consternation for the FAA. Whereas the process for using proprietary data generated from flight data recorders has been well established for investigations, routine maintenance, and flight safety enhancement, there is no equivalent formalized process for including open-source data in an airline’s SMS. While open-source data provide exciting and beneficial opportunities to further safety, policy and guidance around its use are lagging.

Because open-source data inhabits a space outside of formal SMS programs, a number of U.S. airlines have experienced instances in which members of the public have obtained ADS-B data and forwarded the information as a hotline message to local FAA field offices. By congressional mandate, the FAA must investigate hotline calls, necessitating investigation by an FAA inspector. In the best-case scenario, the open-source data would be siphoned into the airlines’ safety management system and employees would be subject to the protections codified in contractual agreements.

However, in the worst-case scenario, open-source data could be either discarded or used to refute other proprietary data sources. For example, ADS-B data could indicate a rejected takeoff (RTO) in a low-speed regime, while the pilot may report the RTO in a high-speed regime. This differentiation has important consequences regarding airworthiness and maintenance costs. During the investigation of such an event, disputes may arise regarding the accuracy or honesty of the pilot’s report if the event review committee were to consider open-source data and compare them to the pilot’s report. Airlines that do not have a mature just culture may risk excluding such a report from their program, to the detriment of overall safety culture.

Partnership extends to national information sharing in the U.S., namely, the Aviation Safety Information Analysis and Sharing Program (ASIAS). Funded by the FAA, ASIAS collects identified flight data and textual safety reports as well as a panoply of other information, both proprietary and open-source. These data are then fused together to obtain a holistic picture of the aviation environment. Because of the potential variation in the quality of open-source data, the robust and multifaceted nature of the ASIAS data streams heightens the veracity of any analysis. The takeaway is that ASIAS has created a robust data architecture for the inclusion of both proprietary and open-source data.

Efforts are being made to create information sharing databases akin to ASIAS outside of the U.S. These are being initiated by vendors of flight data monitoring programs, SMSs, and aviation trade organizations. While it is commendable that this many entities see the value in data sharing, the siloed nature of their scope may limit the extent of their effectiveness.

An altruistic partnership program on a global scale that could share information across physical, legal, and commercial borders would be the most beneficial approach for the industry.

4. The Legal Landscape

Traditional aviation data contained in an airline's SMS are protected by law in the U.S.; however, open-source data are not. Regarding FOQA programs, the law states that "the Administrator will not use an operator's FOQA data or aggregate FOQA data in an enforcement action against that operator or its employees when such FOQA data or aggregate FOQA data is obtained from a FOQA program that is approved by the Administrator" [3]. The extent of this protection includes FOQA programs as approved by the FAA, and stops there.

Open-source aviation data are not protected under U.S. law because they are considered ancillary data source and not a component of a FOQA program. What this means in practice is that the FAA could run a parallel investigation of an event based on open-source data from an aircraft, and not from FDR-based flight data. Prior to the advent of ADS-B, the only data available in the event of a parallel investigation by the FAA were from external ground-based surveillance radars. Conducting an investigation based on open-source airplane-generated data would bypass the FOQA program, yet be perfectly legal. Even more alluring may be the speed at which the data could be obtained. No longer would days or weeks elapse waiting for flight data to enter a company's FOQA program; instead, an investigation could begin in real-time.

Legal questions surrounding open-source flight data exist with respect to the unknown unknowns as well. The FDR today is used for a purpose it was not entirely designed for originally, in that it is now a proactive safety management tool rather than a reactive one. This could be the case for open-source flight data as well. For example, today the flight data recorder and cockpit voice recorder are considered proprietary data; yet, many parameters that are recorded by the FDR and flight management system (FMS) are transmitted in an open-source format via ACARS, including takeoff and landing data, waypoint reporting, and aircraft diagnostics. In the push for single-pilot operations, manufacturers have suggested the need for pilot health monitoring (PHM) programs [5], which have been compared to aircraft engine health monitoring (EHM). Could biometric information be included in routine ACARS transmissions one day? Even if biometric data were kept proprietary, would this be considered "flight data" and be afforded the protections of FOQA data?

Lastly, the ability to see personal flight data on tablets presents a paradox with respect to data and protection. For pilots to buy in to a personal flight data monitoring program, they likely need to believe that the airline cannot monitor them individually. Yet, it is legally unresolved what would happen if an accident were to occur, and the airline held data that indicated a consistent performance deficiency that ultimately led to an accident. For example, if a personal flight data monitoring program showed a pilot landing long over time and this led to a runway excursion, would the pilot's previous data then become discoverable in the investigation? What responsibility does the airline have to identify substandard performance and remedy it before it develops into an accident? The new data world in aviation presents many legal issues that currently have more questions than answers.

5. The Ethical Dimension

A new line of inquiry has emerged in light of the present-day surge of technology, namely, that of data ethics. Considering the proliferation of all forms of data in modern society, scholars have understandably raised concerns with ethical issues surrounding data use. Fields such as statistics and computer science, which have permeated the data revolution, have previously focused exclusively on technical matters; however, the way in

which technology is being developed and used is bringing to light subtle ethical issues in these fields [6].

The use of data in aviation touches every aspect, from aerodynamics to customer satisfaction surveys. Safety management systems are metric-driven, and the drive towards quantification furthers the chasm between numbers and their ethical value. It can be argued that a feature of a mature domain is the creation of and adherence to ethical standards [7]. As of now, the issue of data ethics has been absent in the aviation literature, and it may be beneficial for the industry to address this.

Today's airline data ecosystem resembles a jungle. Proprietary and open-source data are generated by the terabyte. For example, take an engine health monitoring (EHM) program. Engine performance data are sent in real-time to the manufacturer with the aim of identifying problems and preventing downtime operational disruptions. Events have happened in which, unbeknownst to the pilots, out-of-tolerance engine parameters were transmitted to an engine manufacturer while airborne and then sent to the airline. A flurry of activity was undertaken, including what-if engine failure scenarios, airport diversion planning, and emergency response drills. However, the flight crew was the last link in the communication chain, and were not aware of the situation until after landing. Understandably, a potential engine failure is something all pilots should be privy to; however, the complex reality of present-day data flow processes impedes this effort.

Data ethicists state that data generators, in this case pilots, should be provided with a transparent view of how their data are being used [8]. Many pilots may believe that their flight data are merely being captured by the flight data recorder and exclusively used for the internal FOQA program. However, the reality is that their data are moving in all manner of directions, both internal and external. Based on the author's experience, many pilots are shocked to learn that their identified data are fused with other data sources as part of the ASIAs program. The ethical use of data involves, in part, knowing who has access to data and when and how it is used. By this standard, the U.S. industry seems to be coming up short.

As discussed earlier, legal shortcomings pose a challenge to the new reality of the use and protection of flight data. Regulators are clearly being outpaced by technological developments, and the impact of these developments has yet to be determined in the case of flight data [9]. Because of this, subject matter experts are increasingly called upon to anticipate impacts and think proactively about data ethics. This has yet to be addressed in aviation. Moreover, technologies are reshaping the distribution of power and responsibility. For those in aviation flight data programs, this is seen as a shift away from labor unions' strong influence over FOQA programs, and a redistribution of it to individuals via personal access to data. This poses a potential threat, as bottom-up individual sensemaking when viewing personal flight data may become the most salient experience for the pilot, replacing the formalized top-down FOQA process that the industry has relied upon to date.

6. Recommendations

Based on the discussion points raised above, the following recommendations can be generated:

- U.S. airlines should ensure that FOQA gatekeeper pilots are not excluded when designing programs that provide personal flight data to pilots. Measurements, exceedance values, and event sets should continue to be created in a collaborative setting.
- U.S. regulations that protect FOQA data from enforcement action should be expanded to include all sources of flight data, both proprietary (FDR, QAR, EHM, etc.) and non-proprietary (ADS-B, ACARS, etc.).
- ICAO SARPs can be developed to include best practices for use of open-source data.
- Training and education should be provided for flight data analysts to integrate open-source data into their safety management systems.
- A global data repository modeled on ASIAs could be created. Such a system should be agnostic to vendor, manufacturer, or other commercial interests.

- Labor organizations should develop policy language addressing best practices for personal flight data monitoring programs.
- Efforts should be made to educate frontline employees on the various recording devices captured by an aircraft, their transmission, the extent of identification, and the duration of retention.

7. Conclusions

Aviation has evolved to be one of the safest modes of transport, thanks in part to its evolution from a reactive approach towards a proactive safety management approach, and ultimately a predictive approach. This paper has adopted a similar framework; while the issues raised herein may not be manifested in today's aviation data landscape, the rapid expansion of aviation data use necessitates an attempt to stay one step ahead and predict future states.

Pilot labor representatives have worked in harmony with their airline counterparts to create voluntary safety programs that have contributed to the remarkable safety record of U.S. airlines. Pilot gatekeepers should remain essential ingredients in flight data monitoring programs, ensuring that trust and accountability remain front and center.

It is unknown to what extent personal flight data may be used in an accident investigation. Without a robust data protection legal framework that includes both proprietary and open-source data, pilots may be exposed to harm if data are used in ways counter to just culture principles. Finally, big aviation data should be included in the discussion of data ethics.

Open-source data offer many opportunities to enhance the aviation domain, and these are only expanding. Perhaps most exciting is the opportunity for open-source aviation data to be disseminated and studied in the academic setting. Previously, obtaining flight data parameters would have been nearly impossible, as airlines were averse to sharing such data outside of formalized internal programs. Organizations such as the OpenSky Network have created communities that have brought together experts in open-source data, applied standards of computer science, and published high-quality peer-reviewed articles.

In summary, this paper raises concerns over the rise of open-source data in aviation, as well as novel uses of proprietary data. Three topics have been presented for discussion, with analysis undertaken and recommendations posed. Considering the dizzying speed with which the topics discussed in this paper have emerged, it will require significant effort by all stakeholders to take a proactive approach in order to ensure that the future use of flight data is as successful as the past.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The author would like to thank the OpenSky Network for supporting an open and collaborative dialogue between academia and industry.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ACARS	Aircraft communication addressing and reporting system
ADS-B	Automatic dependent surveillance–broadcast
ASIAS	Aviation Safety Information Analysis and Sharing
EHM	Engine health monitoring
FAA	Federal Aviation Administration
FDR	Flight data recorder
FOQA	Flight operations quality assurance
ICAO	International Civil Aviation Organization
PHM	Pilot health monitoring
RTO	Rejected takeoff
SARP	Standards and recommended practices
SMS	Safety management system

References

1. Achmad, W. Citizen and Netizen Society: The Meaning of Social Change From a Technology Point of View. *J. Mantik* **2021**, *5*, 1564–1570.
2. Yermack, D. Tailspotting: Identifying and profiting from CEO vacation trips. *J. Financ. Econ.* **2014**, *113*, 252–269. [[CrossRef](#)]
3. FAA. *Flight Operational Quality Assurance*; U.S. Department of Transportation: Washington, DC, USA, 2004.
4. Mills, R.W. The promise of collaborative voluntary partnerships: Lessons from the Federal Aviation Administration. *Collaborating across Boundaries*; IBM Center for the Business of Government: Washington, DC, USA, 2010.
5. Schmid, D.; Stanton, N.A. Progressing toward airliners' reduced-crew operations: A systematic literature review. *Int. J. Aerosp. Psychol.* **2020**, *30*, 1–24. [[CrossRef](#)]
6. Hand, D.J. Aspects of data ethics in a changing world: Where are we now? *Big Data* **2018**, *6*, 176–190. [[CrossRef](#)] [[PubMed](#)]
7. Martin, C.W.; Vaught, W.; Solomon, R.C. *Ethics across the Professions: A Reader for Professional Ethics*; Oxford University Press: New York, NY, USA, 2010.
8. Herschel, R.; Miori, V.M. Ethics & big data. *Technol. Soc.* **2017**, *49*, 31–36.
9. Vallor, S. An introduction to data ethics. In *Course Module*; Markkula Center for Applied Ethics: Santa Clara, CA, USA, 2018.