International Civil Aviation Organization

WORKING PAPER

STA/11-WP/3
24/3/22
Revision No.1
1/4/22

ELEVENTH SESSION OF THE STATISTICS DIVISION

Virtual, 4 to 8 April 2022

Agenda Item 2: Big data analytics

BIG DATA ANALYTICS

(Presented by the Secretariat)

EXECUTIVE SUMMARY

This working paper reports on the on-going big data activities, including the collaboration with the United Nations (UN) and international organizations, as well as the provision of dashboards to Member States and aviation stakeholders to support their current and post COVID-19 pandemic planning and implementation efforts. It also reports on activities in accordance with the recommendations of the Third Meeting of the Aviation Data and Analysis Panel (ADAP/3).

Action by the Division is indicated in paragraph 6.

References:	Doc 10140, *Assembly Resolutions in Force (as of October 2019)* Annex 9 - *Facilitation* A40-WP/19-EC/5, ICAO Statistics Programme and Big Data Analytics ADAP/3-WP/6, Use of Big Data by ICAO *Report of the Third Meeting of the Aviation Data and Analysis Panel* (ADAP/3, Yellow Cover) *Draft Guidance on the use of big data in aviation* Doc 9944*, Guidelines on Passenger Name Record (PNR) Data.*

1. INTRODUCTION

1.1 The 40th Session of the Assembly requested the Council to: a) continue to explore ways of closer cooperation with the United Nations (UN), its agencies and other international organizations in the collection and distribution of aviation data, statistics and analysis, including big data as required; and b) continue to collect, process and analyze aviation data, including big data, while ensuring the harmonization of aviation data and statistics from different sources in order to facilitate the provision of accurate, reliable and consistent data required for informed decision-making by States.

1.2 Activities carried out on big data, in accordance with the A40 decision, were further reported to the Third Meeting of the Aviation Data and Analysis Panel (ADAP/3) in June 2021. The Panel recommended the Secretariat to determine the feasibility, benefits, and risks of using big data to assist in

contact tracing of passengers during public health emergencies, and lead the development of a guidance document on the use of big data in aviation (Recommendation ADAP/3.5).

2. SOURCES OF BIG DATA

2.1 ICAO has been gradually expanding its activities related to big data, namely: a) Market Intelligence Data Transfer (MIDT); b) Automatic Dependent Surveillance-Broadcast (ADS-B); and c) Commodity and Trade (COMTRADE) and e-commerce transactional data (E-com). These data sets represent an all-encompassing coverage of traffic and operational data at a very high level of granularity.

2.2 The MIDT contains passenger booking data of the Global Distribution Systems (GDS) supplemented with the direct sales of air carriers, providing true origin-destination records of more than four billion passengers and 36 million departures annually (in 2019). The Secretariat verifies, validates and aligns the MIDT data with data reported by States.

2.3 The ADS-B is "a means by which aircraft, aerodrome vehicles and other objects can automatically transmit and/or receive data such as identification, position and additional data, as appropriate, in a broadcast mode via a data link" (Annex 11 – *Air Traffic Services* refers). Approximately 600 million rows of ADS-B data on aircraft positions (departure, en-route and arrival in intervals of one minute) are recorded monthly, covering a vast majority of scheduled passenger and freighter operations, as well as charter, business jets and other commercial operations. The algorithms developed by ICAO verify the ADS-B data and complement the data with additional information of carrier code, aircraft type, Flight Information Region (FIR), type of operation and distance flown while the ADS-B data is being streamed into the Cloud.

2.4 The COMTRADE data provided by the United Nations (UN) contains billions of records of commodity movements, including origin and destination countries, cost, insurance, mode of transport, and the type of transaction. E-com records provided by the Universal Postal Union (UPU) contains e-commerce parcel transactions by date, time, class, weight, number and origin-destination city-pair.

3. COLLABORATION ON BIG DATA ANALYTICS

3.1 The aforementioned big data sets have been collected, processed and analysed in close collaboration with other UN agencies and international organizations such as the International Transport Forum (ITF) at the Organization for Economic Co-operation and Development (OECD), the UPU, the UN Conference on Trade and Development (UNCTAD) and the World Customs Organization (WCO). In addition, internal collaboration within the Secretariat is ongoing to support other activities of the Organization. Highlights of these collaboration are presented below.

3.2 The MIDT data was provided to the World Bank along with the methodology to calculate an air connectivity index by State. The data and methodology is also provided to the International Air Transport Association (IATA) and the Air Transport Action Group (ATAG) to disseminate the air connectivity index and ranking by State through the *Aviation: Benefits Beyond Borders*[1] report.

3.3 The ADS-B along with the metadata was provided to the United States Federal Aviation Administration (FAA), the European Aviation Safety Agency (EASA) and members of the Committee on Aviation Environmental Protection (CAEP), to assist in their validation of the Common Operations

[1] https://aviationbenefits.org/downloads/aviation-benefits-beyond-borders-2020/

Databases (COD). The COD is the main database used by CAEP in the development of fleet forecasts and trend assessments.

3.4 The ICAO Office of Environment has developed a process to fulfil this requirement on Carbon Offsetting and Reduction Scheme for International Aviation (CORSIA) data gap filling utilizing ADS-B data in combination with other datasets and environmental tools. This data gap filling process under CORSIA is carried out by the ICAO Office of Environment annually (per Annex 16, Volume IV).

3.5 ICAO and OECD are collaborating in using ADS-B big data for the purposes of Environmental Economic Accounting. This collaboration is designed as a resource for national statistical systems and to facilitate the development of Air Emission Accounts (AEA) databases, following the System of Environmental-Economic Accounting (SEEA) international standard.

3.6 In coordination with the European Organisation for the Safety of Air Navigation (EUROCONTROL), the ADS-B data has now been integrated with the ICAO Aero-Tariffs Application and made available to Member States and Air Navigation Services Providers (ANSPs) through the iCADS platform (https://data.icao.int/AeroTariffs/).

3.7 With regard to the COMTRADE and E-com data, ICAO and the UPU have embarked on a joint project using both data to identify and quantify logistical constraints on e-commerce activity.

3.8 The above projects were presented to ADAP/3, which strongly supported the continuation of collaboration between ICAO, Member States and other organizations in the area of big data analytics, and suggested that the Secretariat share the results of the projects with Member States.

4. BIG DATA DASHBOARDS

4.1 The Aviation Data and Analysis Section of ICAO worked with the United Nations Global Big Data Group to develop business intelligence dashboards for data driven decision making. The online dashboards have been developed using the state-of-the-art Tableau platform, which provide visualization and querying capabilities.

4.2 The ICAO dashboards cover a wide array of analytics ranging from operations, fleet usage, passenger carried, Flight Information Region (FIR) traffic, and revenue impact on airlines, airports and air navigation services providers (ANSPs) at different levels of granularity. They allow for the visualization and download of data by State, region, route group and FIR. Information contained in the dashboards are continuously updated and can be used to monitor and assess the evolving impact of the pandemic on air transport. The dashboards will be extremely valuable to the ICAO Council and Member States in their current and post COVID-19 pandemic planning and implementation efforts.

4.3 In this regard, the dashboards will be made available to all Member States and a State letter will soon be issued to communicate the access codes and an exclusive licence. States will be invited to provide their comments and feedback on the use of the dashboards, which will in turn be compiled by the Secretariat for further improvement. Access to the dashboards will also be provided to external stakeholders in particular the UN and collaborating international organizations such as World Bank, UPU, UNCTAD, and OECD.

4.4 Moreover, the Secretariat has developed a guidance document on the use of big data in aviation, which also provides technical and practical information on the dashboards (https://www.icao.int/Meetings/STA11/Documents/Dashboard_Document.pdf).

5. BIG DATA FOR PASSENGERS' FLOW ANALYSIS DURING PUBLIC HEALTH EMERGENCIES

5.1 Big data is a key enabler of these monitoring activities, through its ability to deliver connectivity patterns of travellers using information from the GDS. ICAO has been providing structured connectivity data of air travellers to the World Health Organization (WHO) to segment regions and populations at risk due to travel from and to areas impacted by public-health emergencies.

5.2 During ADAP/3, the Panel discussed the possibility to monitor connectivity of international passengers in real time during times of public-health emergencies, using the potential of Big Data and particularly the Central Reservation Systems (CRS) and Passenger Name Record (PNR) data. While noting the interests from States in this aspect, it was also stressed that such work be carried out in adherence to privacy concerns of stakeholders, existing Standards and Recommended Practices (SARPs) enacted under Annex 9 - *Facilitation*, and avoid duplication of efforts. Furthermore, the Panel was invited to note that not all States are using PNR data, and there are significant gaps as some parts of the globe are not covered. In 2022 only 60 States have implemented a PNR data system.

5.3 The Panel recommended that the Secretariat should coordinate with the Facilitation Panel (FALP) in accordance to its Terms of Reference, to establish a multi-disciplinary working group tasked with determining the feasibility, benefits, and risks of using big data to assist in passengers flow analysis during public-health emergencies.

5.4 Following internal coordination, it is noteworthy that Amendment 29 to Annex 9 becoming effective on 18 July 2022, introduces a definition of contact tracing[2] and related new provisions for data collection. Two forms have been updated and developed, namely "Public Health Passenger Locator Form" and the ''Public Health Passenger Self-Declaration Form" that should be made available in digital or paper format in the context of the *No Country Left Behind* initiative. These forms have been defined to collect travel itineraries or contact information of passengers and/or crews as well as information concerning their health status for the purpose of a specific incident management.

5.5 Besides, PNR[3] data may include many separate data elements under a non-uniformed way hence differing from one airline to another. In practice, aircraft operators capture only a limited number of data as key elements for the creation of a PNR, such as all check-in information, all seat information, all baggage information and "go-show" and "no-show" information. Accordingly, the structure of individual PNRs and the amount of data they contain will vary widely. Specific data elements that may be available from an operator's system(s) are set out in the Appendix as defined in Doc 9944, *Guidelines on Passenger Name Record (PNR) Data* and no health data is mentioned.

5.6 Internal coordination with the Aviation Security and Facilitation Branch has commenced to explore the potential future discussions with the relevant meetings of FALP in 2023 and onwards. Updates on the status of such discussion will be provided to the Panel.

---

[2] Contact tracing is the practice of identifying, notifying, and monitoring individuals who may have had close contact with or who have been exposed to, and possibly infected by, a person having a confirmed or probable case of an infectious disease as a means of controlling the spread of infection. The confirmed or potentially infected person's identity is not discussed with contacts, even if asked.

[3] PNR data, in the air transport industry, is the generic name given to records created by aircraft operators or their authorized agents for each journey booked by or on behalf of any passenger. The data are used by operators for their own commercial and operational purposes in providing air transportation services.

6. ACTION BY THE DIVISION

6.1 The Division is invited to:

a) note the work of ADAP in the area of big data analytics;

b) invite Member States to access the big data dashboards and guidance document for their data-driven decision making, planning and implementation needs; and

c) support the on-going collaboration on big data analytics with Member States, UN, international organizations and other stakeholders.

— — — — — — — —

APPENDIX

Passenger Name Record (PNR) DATA ELEMENTS

An operator's system(s) may include the following data elements:

Data groups or categories	*Component data elements*
PNR name details	Passenger name, family name, given name/initial, title, other names on PNR
Address details	Contact address, billing address, emergency contact, email address, mailing address, home address, intended address [in State requiring PNR data transfer]
Contact telephone number(s)	[Telephone details]
Any collected API data	Any collected API data, e.g. name on passport, date of birth, sex, nationality, passport number
Frequent flyer information	Frequent flyer account number and elite level status
PNR locator code	File locater number, booking reference and reservation tracking number
Number of passengers on PNR	[Number]
Passenger travel status	Standby information
All date information	PNR creation date, booking date, reservation date, departure date, arrival date, PNR first travel date, PNR last modification date, ticket issue date, "first intended" travel date, date of first arrival [in State requiring PNR data transfer], late booking date for flight
Split/divided PNR information	Multiple passengers on PNR, other passengers on PNR, other PNR reference, single passenger on booking
All ticketing field information	Date of ticket issue/purchase, selling class of travel, issue city, ticket number, one-way ticket, ticket issue city, automatic fare quote (ATFQ) fields

All travel itinerary for PNR	PNR flight itinerary segments/ports, itinerary history, origin city/board point, destination city, active itinerary segments, cancelled segments, layover days, flown segments, flight information, flight departure date, board point, arrival port, open segments, alternate routing unknown (ARNK) segments, non-air segments, inbound flight connection details, on-carriage information, confirmation status
Form of payment (FOP) information	All FOP (cash, electronic, credit card number and expiry date, prepaid ticket advice (PTA), exchange), details of person/agency paying for ticket, staff rebate codes
All check-in information*	Generally available only after flight close-out: check-in security number, check-in agent I.D., check-in time, check-in status, confirmation status, boarding number, boarding indicator, check-in order
All seat information	Seats requested in advance; actual seats only after flight close-out*
All baggage information*	Generally available from DCS only after flight close-out: number of bags, bag tag number(s), weight of bag(s), all pooled baggage information, head of pool, number of bags in pool, bag carrier code, bag status, bag destination/offload point
Travel agent information	Travel agency details, name, address, contact details, IATA code
Received-from information	Name of person making the booking
Go-show information*	Generally available only after check-in and flight close-out: go-show identifier
No-show information*	Only available after flight close-out: no-show history
General remarks	All information in general remarks section
Free text/code fields in OSI, SSR, SSI, remarks/history	All IATA codes

— END —