International Civil Aviation Organization

WORKING PAPER

THIRTEENTH AIR NAVIGATION CONFERENCE

Montréal, Canada, 9 to 19 October 2018

COMMITTEE A

Agenda Item 1: Air navigation global strategy
1.1: Vision and overview of the sixth edition of the GANP

BIG DATA IN THE CONTEXT OF (ATM) INFORMATION MANAGEMENT AND BEYOND

(Presented by the United Arab Emirates)

EXECUTIVE SUMMARY

This paper aims to stimulate debate on the use of Big Data in air traffic management (ATM). Information management (IM) is gaining a new dimension in the context of the global air navigation system with the integration of aeronautical, flight and flow, meteorological and surveillance information for ATM decision-making processes. In this regard, size, diversity and quickness of access to the information will be determining factors in the optimisation of the current and future ATM applications. This paper provides an introduction to the concept of Big Data, explores its potential use within the context of a global air navigation system and recommends its consideration within the *Global Air Navigation Plan* (GANP, Doc 9750). Also, this paper is proposing a forward-thinking way for future air traffic management (ATM)/aviation database management direction;

Action: The Conference is invited to agree to the recommendation in paragraph 3.4.

1. INTRODUCTION

1.1 Aeronautical information services are evolving towards aeronautical information management, where more digital data is replacing the traditional means of collecting, processing and distributing relevant information.

1.2 The Global Air Navigation Plan (GANP, Doc 9750) and its associated aviation system block upgrades (ASBUs) place aeronautical information, together with flight and flow, meteorological (MET) and surveillance information at the core of air traffic management (ATM) by creating an overarching information management layer around the seven ATM concept components (ref. the *Global Air Traffic Management Operational* Concept (Doc 9854).

1.3 Information is relatively well defined and structured (although evolving conceptually) within each of the domains, however, when gathered together in the ATM context, information becomes significantly heterogeneous in use and applicability.

1.4 The different nature of the information together with, at global scale, the increasing size of the data and the applications/services that require (near) real-time feed and analysis may qualify ATM as a good candidate for the Big Data appliance.

2. DISCUSSION

2.1 With digitization, the advent of different social media, cloud computing, the adoption of hand-held digital devices and, generally, the explosion in the usage of the Internet, a massive amount of data is being generated on a continuous basis. No one can disagree with the fact that the Internet has changed the way businesses operate around the world. Today, the rate of data generation is very high and the type of generated data overdoes the capability of existing data storage practices. This phenomenon is called Big Data.

2.2 Big Data is characterised by the "three Vs":

a) volume (*too big*): organizations must deal with petabyte-scale collections of data;

b) velocity (*too fast*): data is generated and must be analysed and processed quickly (e.g. Facebook: 500 Terabyte/day); and

c) variety (*too hard*): data does not fit smartly into an existing processing tool i.e. increasingly different type of data (heterogeneous) from unstructured (raw files), structured files (relational databases) to social media conversations, photos, sensor data, video and voice recordings.

2.3 Depending on the domain or business using Big Data analysis, two additional "Vs" are frequently considered:

- veracity is the quality of the data in the respective context i.e. not 100% correct; and
- value is the most beneficial aspect i.e. the ability to turn data into value.

2.4 There is a great challenge not only to store and manage the tremendous volume of data, but also to analyse and extract meaningful information from it. The state of the art is that, with respect to data size, the traditional relational database management systems (RDBMSs) actually do satisfactorily well because they are being able to handle multi-petabyte databases.

2.5 On the other hand, new existing platforms are supporting relatively sophisticated analysis, however, they are not designed to scale datasets that exceed even a single machine's memory. In other words, the present technologies lack suitably compliance with the five "Vs" metrics mentioned above.

2.6 Research projects are attempting to "bridge the gap" between large-scale data processing platforms and analysis packages in aiming to address all concerns. By oversimplifying things, the typical approach is either (1) to extend the relational model to cover necessary features for analysing and visualizing data or, (2) to extend capable and well-known software analysis solutions to handle Big Data sizes.

2.7 The need of Big Data analysis primarily came from big companies like Google and Facebook. However, there is an essential similar need within the aviation sector: the availability of large "accumulations" of data that are becoming available due to ICAO Standards and Recommended Practices (SARPs) e.g. State Static Data Operations (SDO), regional SDOs, terrain and obstacles (ETOD) files, airport mapping databases, digital datasets, operational traffic/fast time simulation data, airlines flight dispatch briefings, repetitive filed flight plan (FPLs) and data-house applications (FMS procedures coding, moving maps, synthetic vision, etc.).

2.8 Moreover, the aeronautical information data process chain is similar to the data value chain used to process Big Data. Specifically, the Data Value Chain refers to a set of activities, e.g. generation, collection, transmission, pre-processing, storage to analysis and decision-making, that create value from available data. The final goal is "to create value" i.e. based on data analysis and visualization, the decision-maker can decide whether and how to reward an optimistic behaviour and change a negative one.

2.9 In the same way, the Aeronautical Data Chain is a conceptual representation of the "journey" that aeronautical data is taking from its origination to its end-use. Although aeronautical data chain primarily addresses "AIM actors" i.e. data originators, aeronautical information services (AISs), data integrator/provider and air/ground end-users, in the background, it also includes processes and activities. All "players" are maintaining the data in data quality report (DQR)-wise as well as they are contributing for growing-value of the data i.e. data as originated is elevated in the context (information) and finally, in an interoperable manner, is consumed in an application.

2.10 Furthermore, the adoption of the industry concept "*360-degree view of the customer*" could be valuable. This concept embraces the function of real time analytics using the Internet of Things (IoT) and Big Data. A viable 360-degree view needs to include a scrutiny of past, present and future "customer" information, hence, aviation wise, the past-present and future history of, generically speaking, aeronautical data/information. IoT technology ensures the view on the data "past" + "present" situation. Furthermore, by applying the "big analysis" feature, a predictive (analytical) "future" of the information would have been generated.

2.11 The ATM system is fundamentally based on the cumulative data and information that enable the seven concept components: aerodrome operations, airspace organisation and management, demand/capacity balancing, conflict management, traffic synchronization, airspace users' operations and ATM service delivery management. Each of these components is using data and information from various sources that needs to be stored, combined, visualized, searched, analysed, processed and transported.

2.12 As such, new perspectives and opportunities to "extract more value" may open up in areas that represent reasons for adopting Big Data as follows:

2.12.1 **Stop wasting data exhaust** – Data exhaust occurs usually when the ability to produce data greatly exceeds the ability to store and manage it. The introduction of metadata in applications such as AMDB has the potential to double (or even exceed) the size of the actual data. Similarly, the management of ETOD increases the size and the diversity of data that is used for airspace organisation, instrument procedure design, etc. Big Data may be an answer to better organise, manage and produce derived data based on the above.

2.12.2 **Save time and money -** When data is captured, it is typically stored in disconnected areas (silos). This increases the amount of time it takes to find and produce information. Adopting Big Data can optimize techniques and technologies that can connect data sets together, the access and response time in ATM "mission critical" applications and, at the same time, the data quality can be improved.

2.12.3 **Improve performance** – Big Data enables information from both inside and outside the system to be combined, so that key performance indicators (KPIs) can be developed and acted upon. With the increasing dependency on data, the ATM system needs to develop and to constantly monitor such KPIs (e.g. for the performance-based approach in navigation, communication, surveillance, etc.).

2.12.4 **Improve information services** – Data from a variety of sources can be combined to improve existing services and to create new ones. Upon the system wide information management (SWIM) implementation, information services will heavily rely on accessing data from more and more sources spread over the entire global aviation system. Their ability to interpret and combine such data/information may be significantly enhanced by considering Big Data techniques.

2.12.5 **Improve decision-making** – Big Data is a powerful means for making informed decisions. ATM could also benefit from it in strategic areas e.g. airspace user operations, demand/capacity balancing, etc. as well as in addressing safety and security issues.

2.12.6 **Innovation** – Big Data diminishes the need to rely on preconceived ideas or assumptions, by enabling innovators to analyse and experiment using real data, in real time or even in fast time simulations. Analytics techniques such as machine learning algorithms, classification tree analysis, social network investigation can be combined with Big Data technologies specific to very large data sets in order to achieve innovations that may be difficult (even impossible) in different circumstances.

2.12.7 **Governance** - One key aspect of employing Big Data is represented by corporate governance. The effective governance of the global ATM system at the macro-level will be the prerogative of ICAO. Big Data exercises should be closely monitored, supervised and supported by a high-level Stakeholder i.e. an authority for defining policy, providing leadership, and managing the procedures and resources of the future worldwide ATM system.

3. CONCLUSION

3.1 We have entered the age of "Big Data" which is changing the world! Overgeneralizing, Big Data foundation technology, tools and platforms are there, but so far, a complete "solution" bridging the gap between huge scale data processing and analysis is missing. The most exciting element of Big Data is the rethinking of data management principles for providing "Big Analysis" choice leading to "Big Opportunities" to facilitate the decision-making process.

3.2 Most benefit comes from the opportunity of big data (predictive) analytics i.e. examining large, real-time/archived, structured/unstructured, heterogeneous (audio/video) digital data. A global ATM information management architecture can naturally evolve to include information of all aviation activities (civil, military, industry) i.e. an information management architecture model where large unstructured, semi-structured and structured data are generated and gathered from different sources. As a general principle, ATM/Aviation Big Data should be made available and shared by all systems and actors that comprise ATM in its most generic sense.

3.3 The proposed approach is to have a de-centralized, distributive, and participative solution. Possible consideration to the proposed analogy between the Aeronautical Data Chain processes versus the Data Value Chain model. Governance of aviation Big Data: high profile stakeholders "authority" would agree to formalize and institute a set of protocols for managing, harmonizing and governing the "system".

3.4 The Conference is invited to agree on the following recommendation:

Recommendation 1.1/x - Vision and overview of the sixth edition of the GANP

That the Conference:

a) recognize Big Data as an enabler to optimize the decision-making process in the provision of future air traffic services (ATS);

b) recognize the momentum for a strategic transformation of the aviation sector due to emerging internet-based technologies; and

c) consider Big Data in future developments of the *Global Air Navigation Plan* (GANP, Doc 9750) and specifically in the system wide information management (SWIM) thread.

— END —