Open Data for Air Transport Research: Dream or Reality?

Marc Bourgois and Michael Sfyroeras
EUROCONTROL Experimental Centre (Paris, FR)
{marc.bourgois, michael.sfyroeras}@eurocontrol.int

ABSTRACT

The role of open data in air transport research is analyzed by means of a sample of over 300 research articles. The most used (or available) data types, their sources and their access policies are identified, both for the US and the EU. The analyses show that 70% of research in air transport is heavily reliant on data, that 70% of the data sources are curated by governmental bodies and that the US publicizes a wider set of sources, leading to wider usage. Areas for improving accessibility of (mainly European) data sources are outlined and alternative avenues to obtain data are sketched. The fact that Europe is lagging considerably in making its sources readily available to the research community means Europe missing out on entrepreneurship, innovation and scientific discovery, the presumed benefits of open data.

Keywords

Open Data, Data Policy, Air Transport, Air Traffic Management.

1. INTRODUCTION

This paper is about the role of open data in research, more specifically research in air transport and air traffic management. Air Transport [AT] includes air traffic management, which is the engineering discipline concerned with maintaining a safe and efficient flow of aircraft in the airspace; air traffic management in turn includes the well-known air traffic control activities. Throughout this paper the more general term air transport will be used.

Research in the AT domain is mainly undertaken at governmental research centers such as NASA's Airspace System Program [ASP][1] or the German Aerospace Center [DLR][2] and at aeronautics or transport departments at larger engineering universities. In addition, industrial research departments explore new product opportunities and corporatized air navigation service providers (the organizations responsible for air traffic control) explore new operational concepts through prototyping, simulation and other validation exercises.

The AT research community is not huge, but by no means negligible. An estimated 3000 researchers from 110 entities are currently working on the European ATM modernization programme [SESAR] [3]. This number should be considered a lower bound, as it does not include much of the academic or governmental research capacity. Numbers for the AT research capacity in America are at least that high.

OpenSym '14, August 27 - 29 2014, Berlin, Germany

ACM 978-1-4503-3016-9/14/08…$15.00
http://dx.doi.org/10.1145/2641580.2641602

From our experience in managing European research programmes in AT over the last 10 years, two hypotheses had been taking shape and lead to the analyses reported in this paper: (1) is research in AT becoming more reliant on data and (2) do European authors turn to American sources for data because they are more accessible? Answers to these questions could guide future enabling actions for AT research capacity in Europe.

The remainder of this paper is structured as follows. A motivation for an open data approach in AT research is given in section 2. In section 3 we look at the data types required for AT research and the frequency of their usage. Thereafter we look at the data sources, their current availability and degree of openness, both in the US and EU (section 4). Finally, we recommend a series of specific improvements, especially on the European front (section 5).

2. ROLE OF DATA IN RESEARCH

Research in science and engineering is expected to follow the Scientific Method [4]. An important corollary to the Scientific Method is 'the expectation to document, archive and share all data so that it is available for careful scrutiny by other scientists, giving them the opportunity to verify the results by attempting to reproduce them. The latter practice, called Full Disclosure, also allows statistical measures of the reliability of the data to be established' [5]. From these principles follows the need for access to and openness of data, for any research undertaking in AT. However, the discussion in this paper goes beyond Full Disclosure, beyond making data from past research available for reuse. The greater challenge lies in opening up as many data sources as possible to stimulate novel research, to enable previously unexplored avenues of investigation.

For any research community there are numerous and important benefits in sharing data:

- experiments reusing the same data become comparable and thus significantly strengthening research results;
- time and effort spent in acquiring cleaned and sizeable datasets is reduced, if not eliminated, thus drastically increasing research productivity;
- researchers from other disciplines are constantly looking for unrelated data sets to validate new theories or methods, thus enabling knowledge transfer and serendipitous collaboration.

The enumerated benefits are not theoretical, but manifest themselves in today's research practice. The PLOS ONE journal [6], a peer-reviewed, open-access on-line science journal, already expects Full Disclosure to facilitate independent verification of research results. In the human sciences, the Quality of Government Institute at the University of Gothenburg [7] is famous for the high take-up, by the international research community, of the large cross-national time-series it makes publicly available. Even in the AT domain we have welcomed researchers from unrelated disciplines, such as statistical physics, motivated by the availability of interesting data sets.

In theory, sharing data therefore leads to more research results being produced and more researchers becoming active in the AT domain at virtually no additional costs.

3. WEALTH OF AIR TRANSPORT DATA

Air transport involves many actors, including airlines, booking agencies, airports, ground services, aircraft maintenance, weather services, air traffic control and regulators. Each of these actors collects or generates data to plan, coordinate and analyze its activities. In principle, researchers might want to access any of these data sets to study the performance of any actor's processes.

3.1 Sampling Research Articles

In order to get a grip on what kind of data is currently being used in AT research, we decided to scan a broad sample of readily available research articles. Moreover, in order to be able to study the evolution of the type and frequency of use of the data, the sample needed to be balanced between recent as well as older sets of articles.

As in most engineering domains, conference proceedings take on a dominant role in the dissemination of research results. The most respected conference in the domain is the biennial ATM R&D Seminar [ATMR&D][8] organized by the FAA and EUROCONTROL. This research conference alternates with the International Conference on Research in Air Transportation [ICRAT][9]. In addition, the ICNS and DASC conferences, focusing on enabling technologies in communications, navigation, surveillance and avionics, generally also have tracks on ATM. Although these major conferences are recognized by the appropriate professional associations [AIAA, IEEE], academics also publish in journals in order to secure high impact factors for their research output. Given the relatively small size of the AT domain, no specialized journal dominates the scene. Journal articles written by AT researchers are therefore scattered over a large number of specialized journals, covering psychology, transport, control theory, optimization, automation and many other disciplines. Because no single (or small set of) journals provides a sufficiently large sample, we decided to limit the sample to the largest specialized conferences, namely ATMR&D and ICRAT.

3.2 Generic Data Types

The sample used in this paper consists of all 70 articles of ATMR&D 2013, all 67 of ATMR&D 2005, all 57 of ICRAT 2012 and all 50 of ICRAT 2004. For each of these articles the data type and source were extracted. During this process, a taxonomy of generic data types was built up.

For example, the generic data type *Traffic* refers to any or all of the following elementary data types: origin and destination airport, scheduled timings (departure, arrival, off-block, on-block), planned route, actual position reports (latitude, longitude and altitude, with time stamp). The details of the routes are in turn part of the generic *Airspace* data type, together with air traffic control sector boundaries, opening times and declared capacities, amongst others. The generic *Weather* data type covers items such as forecast and actual wind speed (at different heights), precipitation or cloud formation (type and position); whereas the generic *Environment* data type essentially refers to pollution and noise measurements or estimates. This small selection of data types clearly illustrates the wide scope of data being used in AT research. We argue, from experience in managing European research programmes in AT, that this wealth of data is a major factor in attracting researchers from other disciplines.

Not only is there a wide of variety of data types, there are also large volumes of data involved, making quantitative research feasible. We are not concerned by the volume of data and associated issues in the remainder of this paper, but as an illustration [10][11][12]: there are roughly 10 million flights in Europe every year, transporting 850 million passengers, connecting 100 large to medium airports. Numbers for North America are roughly 30% higher, and the industry grows 2-4% annually in America and Europe.

Tables 1a and 1b: Data Type per Conference (from high to low, in %).

Conference / Data Type	2012 ICRAT	2013 ATMR&D	Total		2004 ICRAT	2005 ATMR&D	Total
Traffic	21	30	26	Traffic	14	25	21
Airport	13	14	14	No Data	40	10	20
Aircraft Performance	12	14	13	Generated Data	8	15	13
No Data	13	8	10	Airport	8	11	10
Weather	5	9	7	Other	8	10	10
Safety	10	3	6	Weather	2	10	7
Generated Data	6	6	6	Aircraft Performance	6	6	6
Airline	6	1	4	Airspace	8	1	3
Other	5	3	4	Safety	0	4	3
Passenger Connectivity	2	4	3	Delay	0	3	2
Delay	3	2	3	ANSP	0	2	1
Airspace	2	2	2	Environment	0	2	1
Human Performance	1	1	1	Passenger Connectivity	3	0	1
Environment	0	1	1	Sectorisation	2	0	1
ANSP	1	0	0	Human Performance	2	0	1
Sectorisation	0	0	0	Airline	0	1	1

3.3 Data Type Usage

Tables 1a and 1b list for each generic data type, the percentage of articles using that data type, ordered from high to low percentage. For completeness we have included two special rows: in row *No Data* are the articles not using any data input at all, and in row *Generated Data* are the articles that use generated (or synthetic) data rather than data extracted from existing databases.

First, consider the total percentages for the most recent conferences in table 1a on the left. Roughly 70% of the research articles report on the use of data from existing databases.

Traffic data clearly corresponds to the greatest need. This may partly be due to the fact that this generic data type aggregates many elementary data types, but because these are often jointly generated, stored and analyzed we do not consider this an anomaly. Other much used data sets are *Airport* and *Aircraft Performance*, and to a lesser extent *Weather* and *Safety*. For all other generic data types our sample is too small to meaningfully distinguish between their respective usages. For experts in the domain, this top 5 does not come as a surprise[1].

Second, consider the evolution in data usage between the most recent conferences in table 1a and the oldest conference editions in table 1b on the right. Overall, there is roughly a 10% growth in the number of research articles reporting on the use of data from existing databases, mainly because articles using *No Data* at all or only *Generated Data* have halved in the 10 year period.

We conclude that research in AT is heavily data-driven and becoming ever more so.

4. CONTRASTING ATTITUDES

Having identified the data types used (or needed) in AT research, we now turn our attention to the sources of the data. Table 2 lists the 20 most widely used sources in our sample of research articles.[2]

4.1 Major data sources

The sources are databases which generally federate data from different numerous data providers. For example, the *Base of Aircraft Data [BADA]* [1] merges aircraft performance data from major aircraft manufacturers and airlines, covering 90% of the current aircraft and operation types in the European airspace. As another example, the *National Offload Program [NOP]* [13] collects flight tracks (flight number, status and position reports) from terminal radar approach control facilities throughout the US. Table 2 consists of a row for each source, with all generic data types (as identified in section 3.2) in the columns. An *x* indicates the data types effectively used by at least one article from our sample, whereas an *o* indicates the data additionally available for use from that source.

The major data sources are in the hands of just a few governmental bodies[3]. On the American side, these are the Federal Aviation Agency [FAA] and the US Government, in particular the Bureau of Transportation Statistics, Department of Transport. On the European side this is essentially the EUROCONTROL Agency [ECTL]. Governments and their Agencies collect data in their role as air navigation service provider, as economic, environmental or safety regulator and in the interest of consumer (passenger) protection.

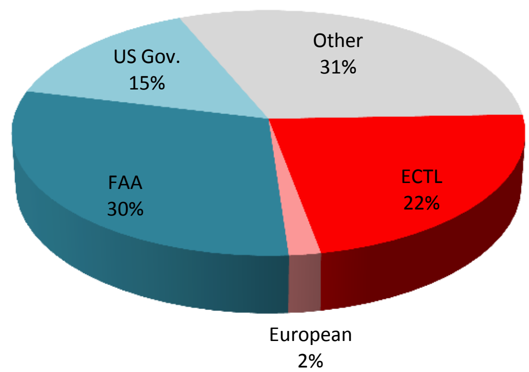

Figure 1: Curators for the 20 major data sources.

Figure 1 classifies the curators of the major data sources. American data sources are twice as often used as European sources. Individual airports, airlines or other data providers make up the remaining 30%. The large difference in usage of American versus European sources is not explained by a larger number of American articles[4], quite on the contrary: in our sample the European articles make up a slight majority. The difference is not explained either by a large number of European articles using American sources, though some minor effect might be present. The explanation seems to lie in more frequent usage of data and more rigorous referencing of the sources by the American articles.

From the above we can tentatively conclude that a wider set of available data from a wider set of sources leads to more frequent usage.

4.2 American Data Policy

The large number of databases made available by the US Government stands out in table 2. Most importantly, all of the data is made available to the public at large, without any significant limitation (as indicated by PA = Publicly Available in the penultimate column of table 2). Even non-American users have open access; this is markedly different from the European situation, as we will see later on.

For several sources, American air transport actors are mandated to report their data to the federating database. For example, airlines are required to file their on-time performance, delay and cancellation data with the Bureau of Transportation Statistics' *Airline Service Quality Performance System [ASQP]* [2] database for all flights accounting for 1% or more of domestic scheduled passengers. As another example, data on passenger connectivity, a

[1] The only anomaly seems to be the low frequency of safety data, which is surprising given the overriding importance of safety in AT; we suspect this may be due to the existence of specialized safety conferences which are not (yet) included in our sample.

[2] In fact the table is constructed from an evolution of our sample, comprising more conferences and adding up to a total of 312 articles.

[3] In the European case these are actually intergovernmental or supranational organizations.

[4] By 'American (European) articles' we mean articles written by authors affiliated to American (European) research entities.

key data set in air transport research as it allows to measure overall system performance from a passenger viewpoint, is collected from a 10% sample of airline tickets in the *Airline Origin and Destination Survey [DB1B]* [3]. The sampling takes care of the legitimate commercial concerns of competing airlines. Given the large annual volume of tickets, even a 10% sample provides a statistically meaningful picture of passenger movements and transfers at the macroscopic scale.

Openness of data reflects the American culture of transparent government manifested in the Freedom of Information Act [FOIA] [14], a piece of legislation obliging the US Government and its agencies to release data unless it falls under one of nine exemptions. The exemptions essentially cater for security, privacy and commercial concerns. President Obama's executive order on 'Making Open and Machine Readable the New Default for Government Information' [15], which updates and strengthens the FOIA, clearly identifies open data as a driver for entrepreneurship, innovation and scientific discovery.

In contrast, the databases provided by the FAA have more restrictive access policies. The popular *Aviation System Performance Metrics [ASPM]* [4] database, focussing on airport analysis, requires registration, but access is readily granted for research purposes (as indicated by UR = Upon Request in table 2). The *Enhanced Traffic Management System [ETMS]* [5] database, focussing on traffic flow predictions, requires proof of a legitimate need (as indicated by UC = Under Conditions). The other FAA databases have restricted access, which is decided on an ad-hoc basis (as indicated by RA = Restricted Access).

In spite of a lack of complete openness, the FAA centrally publicizes and documents its sources: descriptions are up to date, data types are well defined, access conditions are publicly identified and when access is granted, data extraction is straightforward through user-friendly web interfaces.

4.3 European Data Policy

For the purposes of our discussion, the role of the EUROCONTROL Agency in Europe is comparable to the role of the FAA in the US. EUROCONTROL sources provide a wealth of data on traffic forecasts, airspace data, flight tracks and delays *[DDR, PRISME, NM]* [6][7][8], comparable to what is provided through the *ETMS, ASDI* and *NOP* sources on the American side. By far the most popular EUROCONTROL source is *BADA* (section 4.1), which is unique and widely used in Europe as well as America. Access to EUROCONTROL sources is always subject to screening, often requiring proof of legitimate need. Requests 'for research purposes' are generally accepted as corresponding to a legitimate need.

Until very recently, EUROCONTROL had a policy of 'not disclosing data unless'. In 2012, that policy was reverted into 'disclosing unless not' [16]. Though this may seem a dramatic change, and a change for the better, in practice only small steps have been made so far towards opening up the EUROCONTROL data sources. According to an internal report on data requests for the year 2013 [17], requests from outside of Europe, from individuals or not issued through a competent authority are systematically refused. It has to be recognized that the Agency is bound by numerous legacy agreements with the data providers, restricting the use of ATM data, but believes that 'over time the situation will improve' and that these legacy agreements should be screened, and where needed adapted to 'ensure that data received can be used for any purpose set out in the EUROCONTROL convention, beyond what may be laid down in respective agreements'. Note that the EUROCONTROL convention recognizes research co-ordination as part of the Agency's mission.

More urgent for the research community are the observations that: data descriptions for the EUROCONTROL databases are mostly out of date, access conditions remain opaque and data extraction is, with the exception of *DDR* and the Network Manager's business-to business web interface [9], subject to time consuming clarifications and to availability and intervention from EUROCONTROL experts. Not only are the currently used sources and associated tools badly publicized, a recent internal inventory showed that they make up only a fraction of the sources that are used by EUROCONTROL. Numerous sensor monitoring databases (including avionics), phraseology databases, airport performance data, flight cancellations, route charges etc. are all waiting to be discovered and bring benefits to the research community at large.

An important leap forward is expected with the full implementation of the System Wide Information Management concept [SWIM] [10] in which EUROCONTROL has heavily invested. SWIM consists of a set of standards, infrastructure and governance enabling the management of AT information and its exchange between qualified parties via interoperable services. Although this will not necessarily secure open access for researchers, one can anticipate advancements in data definition, organization and exposure with less meandering for access.

The pendant of the US Government on the European side is the Directorate-General for Mobility and Transport of the European Commission and its agencies [EU], in particular the regulatory European Aviation Safety Agency [EASA]. In sharp contrast to the US Government, these EU sources[5] do not appear at all in the list of 20 most used databases (table 2). Although the EU has established rights for public access to documents and legislation on the reuse of public sector information [18], the relevant datasets available on the EU Open Data Portal are few and correspond to aggregated, static snapshots [11]. Open access for scientific publications and research data is also only now being piloted in several of the research programmes of the European Union [20]. One notable exception being the development of an EU spatial data infrastructure, enabling open access to environmental spatial information from legally mandated organisations. This initiative is backed by the INSPIRE directive [12] and full implementation is required by 2019.

4.4 Work-Arounds

The considerable difficulties encountered in obtaining access to the data they require, leads researchers to consider alternative avenues.

Quite surprisingly, some data sets which are being protected for commercial reasons can actually be bought. Although airlines in Europe object to sharing passenger connectivity data, because they claim it might reveal their operational and commercial strategies, that data can be accessed through passenger reservation system providers such as Galileo, Amadeus or Sabre. Unfortunately the cost of procuring access to these commercial databases is prohibitively expensive, but a consolidated data set

[5] Data provision by any of the 28 member state governments and associated agencies has not been included in our discussion, though they are not believed to be substantially different from the over-arching European Union approach.

can be obtained from the IATA PaxIS at a significantly lower price.

Admittedly AT is full of, in principle sound, reasons for not opening up its data: AT has a historic record of terrorist vulnerability and thus needs to cater for security concerns; it also suffers from cut-throat competition and thus needs to heed commercial interests; ATC is highly unionized and thus needs to tread carefully in the minefield of personnel/privacy issues. Nevertheless, the US has largely been able to transcend these counter-arguments, essentially through sophisticated regulation (see the case of [DB1B] in section 4.2) making it difficult to argue why Europe could not follow suit.

Besides the option of buying data, there is an option of trying to obtain the data through non-official sources. We would call this the guerilla approach to data acquisition. In the AT domain it started out with citizens/neighbors of busy airports being skeptical about officially released data and setting up their own networks of low-cost noise measurement sensors [21]. Another transport example concerns traffic jam measurements in urban areas: whereas larger agglomerations might provide web applications displaying road traffic flow data obtained through embedded induction loops, the Google Map application that tracks mobile Android phones is more accurate and definitely lower cost.

Back to AT, the FlightRadar24 [22] application picks up the unencrypted ADS-B signal from commercial aircraft and can thereby reconstruct flight trajectories, a data set most coveted by researchers and notoriously difficult to come by. Analogously to the road traffic jam approach, an Android phone suddenly jumping 800 km within one hour[6] and located at a European airport shortly before and after the jump, would be a sure sign of passenger movement by plane, and could therefor also be used to reconstruct passenger movement and connectivity databases. These guerilla approaches to data acquisition may be less accurate in the case of the noise sensor network or FlightRadar24, the data may nevertheless still be good enough for most research purposes.

5. CONCLUSIONS

In summary, from a sizeable sample of research articles in the AT domain we have identified the major data types currently in use or potentially available, the trends in the usage of that data over a 10 year period, the sources for that data and the widely differing access policies for the US and the EU.

The analyses in the preceding sections show that 70% of research in air transport is heavily reliant on data, that 70% of the data sources are curated by governmental bodies and that the US publicizes a wider set of sources, leading to wider usage.

In order to reach the full promise of open data in AT research, as advocated by both US and EU legislation, we recommend focusing on:

- providing clearer data descriptions, preferably following best practice in Linked Data [23]. A shared ontology of AT data types would dramatically increase the usability of the sources and facilitate ad hoc federations of data across data sources.
- providing a single entry point and self-service web interface to the EUROCONTROL data sources, minimizing human intervention for access validation and data extraction, as it is being piloted by the FAA.
- populating the EU Open Data Portal with more, and more useful AT databases by extending EU regulations on mandatory reporting with open access to raw data, while respecting legitimate security, privacy and commercial concerns, as it is being piloted by the US Government.

Investing in these priorities will lead to increased cross-validation of research results, give preference to fact-based research and lower entry into the AT research domain for experts from other disciplines. Overall, these actions are expected to have an immediate positive impact on research quality and quantity at a low cost.

The evidence and arguments provided in this paper are subject to several caveats: The sample of research articles is still somewhat limited and may not cover all aspects of the AT domain proportionally. Our understanding of the accessibility of and regulations pertaining to the sources is also limited and may require further discussion with the [curating organizations]. Overcoming these shortcomings is subject to on-going work.

6. ACKNOWLEDGMENTS

We thank Stelios Adamopoulos, Paul Bosman and his team (all EUROCONTROL), Andrew Cook (University of Westminster) and José Javier Ramasco (University of the Baleares) for clarifications provided.

7. REFERENCES

[1] NASA: Airspace System Program. DOI= http://www.aeronautics.nasa.gov/programs_asp.htm

[2] German Aerospace Center (Deutsches Zentrum für Luft- und Raumfahrt). DOI= http://www.dlr.de/

[3] Borghini, C: From Innovation to Solution. Presentation at *Helitech International Expo & Conference*. (London, September 24, 2013)

[4] Sherri L Jackson: *Research Methods and Statistics: A Critical Thinking Approach.* Thomson Wadsworth (Belmont, 2008)

[5] Wikipedia, the Free Encyclopedia. DOI= http://www.wikipedia.org/

[6] PLOS ONE Journal. DOI= http://www.plosone.org/static/policies#sharing

[7] Quality of Government Institute, University of Gothenburg: DOI= http://www.qog.pol.gu.se/data/

[8] USA/Europe Air Traffic Management Research and Development Seminars. DOI= http://www.atmseminar.org/index.cfm

[9] International Conference on Research in Air Transportation. DOI= http://www.icrat.org/icrat/index.cfm

[10] EUROCONTROL: 2012 Annual Report (July 2013). DOI= http://www.eurocontrol.int/publications/eurocontrol-annual-report-2012

[11] EUROCONTROL: 2012 Comparison of ATM-related Operational Performance: US/Europe (May 2012). DOI= http://www.eurocontrol.int/publications/2012-US-europe-comparison-air-traffic-management-related-operational-performance

[6] 800 km and one hour roughly corresponds to the average length and duration of an intra-European flight

[12] Boeing: Current Market Outlook 2013 – 2032. DOI= http://www.boeing.com/boeing/commercial/cmo/

[13] J. DeArmon, A. Mahashabde, W. Bade, and P. Kuzminski: Applications of a Terminal Area Flight Path Library (2011)

[14] US Government: The Freedom of Information *Act. 5 U.S.C § 552, As Amended By Public Law No. 104-231, 110 Stat. 3048*

[15] US Government: Making Open and Machine Readable the New Default for Government Information. *Executive Order 13642* (May 9, 2013).

[16] EUROCONTROL: Agency ATM Data Policy and Regulation on ATM Data Protection and Public Access. *Note de Service / Office Note No 18/12* (May 6, 2012).

[17] EUROCONTROL Data Policy Advisory Group: 2013 Report on ATM Data Management. *Internal memo* (January 20, 2014).

[18] EUROCONTROL: EUROCONTROL Data for WP-E Projects. *Internal Note* (July 2014).

[19] European Commission: Legislation on Reuse of Public Sector Information. DOI= https://ec.europa.eu/digital-agenda/en/legislative-measures

[20] European Commission: Guidelines on Open Access to Scientific Publications and Research Data (December 2013). DOI= http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

[21] Deutsche Fluglärmdienst: Messwerte. DOI= http://www.DFLD.de

[22] Flightradar24: Live flight tracker. DOI= http://www.Flightradar24.com

[23] Hall, W: Linked data: The quiet Revolution. *ERCIM News* (January 2014).

8. Selected Data Sources

[1] EUROCONTROL: Base of Aircraft Data. DOI= https://www.eurocontrol.int/services/bada

[2] US Government: Airline Service Quality Performance. DOI= http://aspmhelp.faa.gov/index.php/Airline_Service_Quality_Performance_(ASQP)

[3] US Government: Airline Origin and Destination Survey. DOI= http://www.transtats.bts.gov

[4] Federal Aviation Administration: Aviation System Performance Metrics. DOI= https://aspm.faa.gov/

[5] Federal Aviation Administration: Enhanced Traffic Management System. DOI= http://hf.tc.faa.gov/projects/etms.htm

[6] EUROCONTROL: Demand Data Repository. DOI= https://www.eurocontrol.int/ddr

[7] EUROCONTROL: PRISME. DOI= https://www.eurocontrol.int/prisme/

[8] EUROCONTROL: Network Manager. http://www.eurocontrol.int/network-manager

[9] EUROCONTROL: Network Manager's Business-to-Business interface. DOI= http://www.eurocontrol.int/services/nm-b2b

[10] EUROCONTROL: System Wide Information Management. DOI= http://www.eurocontrol.int/swim

[11] EU Open Data Portal. DOI= https://open-data.europa.eu/en/data

[12] European Commission: Inspire Directive (May 2007). DOI= http://inspire.ec.europa.eu/

Table 2. Data Types for Top 20 Data Sources (from most to least used).

Source	Curator	Aircraft Performance	Airline	Airport	Airspace	ANSP	Delay	Environment	Generated data	Human Performance	No Data	Other	Passenger Connectivity	Safety	Sectorisation	Traffic	Weather	Accessibility	Total
n/a	-		x		x				x	x	x	x				x		-	119
Airport	**various**	x		x	x		x							x		x	x	-	35
Aviation System Performance Metrics [ASPM]	**FAA**			x			x		x			x				x	x	UR	32
Base of Aircraft Data [BADA]	**ECTL**	x																UC	28
No further info	-	x		x	x		x					x		x		x	x	-	21
Enhanced Traffic Management System [ETMS]	**FAA**	x		x					x							x	x	UC	14
Aircraft Situation Display to Industry [ASDI]	**FAA**			x												x		RA	10
Airline	**various**	x	x				x						x			x		-	10
Airline On-Time Performance Data	**US Gov.**		x	x			x					o				x		PA	7
Airport Surface Detection Equipment	**various**			x												x		-	7
Aircraft Engine Emissions Databank	**ICAO**	x						x										PA	6
Network Manager [NM][7]	**ECTL**				x										x	x		RA	6
Demand Data Repository [DDR][8]	**ECTL**						x									x		UC	6
PRISME[8]	**ECTL**			x			x	o								x		RA	6
ANSP	**various**			x		x										x		-	6
Airline Origin and Destination Survey [DB1B]	**US Gov.**	o	o									x	x			x		PA	5
T-100[8]	**US Gov.**	o	o	x									x			x		PA	5
Form 41[8]	**US Gov.**	x	x	x				o				o	o	o		x		PA	5
Rapid Update Cycle [RUC]	**US Gov.**																x	PA	5
Airline Service Quality Performance [ASQP]	**US Gov.**		o	x			x					x				x	o	PA	5
National Traffic Management Log [NTML]	**FAA**			x												x		RA	4
National Offload Program [NOP]	**FAA**															x		RA	4
	Total	40	8	45	4	1	11	1	41	1	70	10	4	4	1	88	18	18	346

x = available data which has been used, **o** = available data which has not been used.
PA = Publicly Available, **UR** = Upon Request, **UC** = Under Conditions, **RA** = Restricted Access

[7] *NM* previously known as *Central Flow Management Unit*

[8] *DDR* is understood to be part of *PRISME*, but referenced independently. Analogously for *T-100* and *Form 41*.