NASA Ames Research Center
IBM Watson Flight Operations Advisor

16 December 2015

This document proposes a Watson solution to address challenges to decision making processes within an airline operations center (AOC). It is part of a larger and longer term engagement with NASA to establish the utility of Watson in a wide variety of use cases and scenarios for aviation. There is also a technology strategy element pertaining to how Watson functionality can be enhanced to support large scale decision making systems that this use case is particularly well suited to drive.

Background

An airline operations center (AOC) is the primary decision-making entity for commercial airlines. Each is staffed by people working in roughly twenty to thirty roles who collaborate to continuously operate an airline's global control process. Dispatchers, who play a central role in this process, oversee the origination, execution, and termination of flights, and are the primary focus of the initial phase of this project.

According to the Code of Federal Regulations, Title 14, from the Federal Aviation Administration:

> *The dispatcher shares joint responsibility with the pilot in command for the planning and operation of a flight. The dispatcher is responsible for delaying or cancelling a flight if necessary.*

In making such decisions, the dispatcher must interface with colleagues (actors) in the other roles in the AOC in order to obtain a complete situational picture and access all relevant data.

The AOC also works closely with entities outside its boundaries, most notably the air traffic control (ATC) system operated by the Federal Aviation Administration (FAA), the airline's own maintenance organization, and several other outside entities including the National Oceanographic and Atmospheric Administration, the National Weather Service, and individual airports.

Dispatchers, and the AOC as a whole, rely on a very large collection of structured and unstructured data and information, at rest and streaming, originating from many sources both within and outside the AOC. Despite employing an array of sophisticated tools and systems to bring relevant information within view of decision makers, the process remains arduous and time consuming, often requiring manual searches, especially in the case of unstructured text documentation and similar sources.

Watson is the first cognitive computing technology platform in the industry. Cognitive systems gain knowledge of the world in the way that humans do, understanding and interacting with humans using natural language. Watson can assist dispatchers and the

AOC as a whole by expanding the range of available information sources and surfacing the most relevant information far more quickly than is possible using keyword or other combinations of manual search techniques.

A team from IBM's Watson Group has been engaged to assist the NASA Ames Research Center in evaluating ways in which cognitive computing in general, and IBM Watson in particular, can assist AOCs to make decision-making processes more efficient and effective. This report documents the results of that evaluation, and presents a draft plan for a phased development of a Watson system for decision support in an AOC that we are calling the Watson Flight Operations Advisor (WFAO).

The AOC Today

The AOC is comprised of numerous positions (sometimes called "desks") where the main actor is a dispatcher or another primary actor (who sometimes is also an experienced dispatcher). The interactions between a dispatcher and a subset of other actors in the AOC are depicted graphically in Figure 1. For the purposes of the initial phase of this project we will focus mainly on the dispatcher, who has authority over decision making during three key phases of a flight.

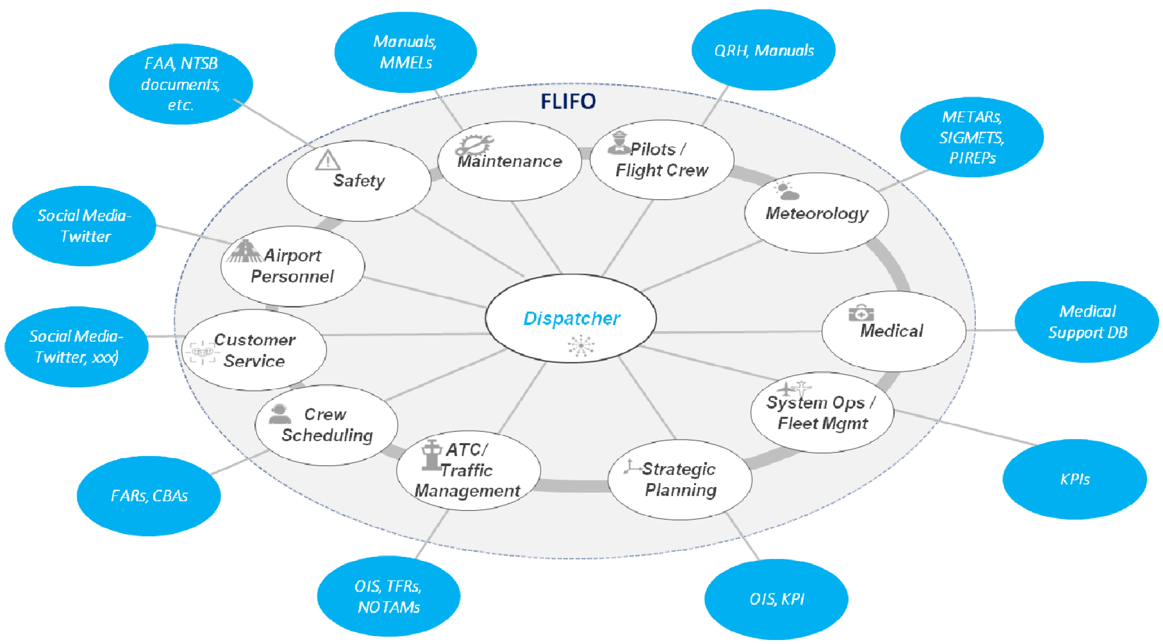

Figure 1: Dispatcher Interaction with Other AOC Personnel and Data Sources (FLIFO = Flight Following)

The first phase is pre-flight planning (route, altitude, speed, fuel planning, and alternate airport selection), during which dispatchers must make decisions affecting delays, cancellations, re-tasking, and other such changes.

During the en-route phase of a flight, the dispatcher is the single point of contact with the airline for the air crews, and oversees possible changes to the flight plan, such as selecting an alternate airport or route, which, in turn, becomes a request to the ATC system by the pilot, and may also include a dispatcher interacting with the AOC's ATC interface desk.

The dispatcher also has authority over flight termination and re-dispatch, communicating with maintenance, and managing delays and other ripple effects. If an aircraft lands at an alternate airport, the dispatcher is responsible for determining safe and economical options to return the aircraft and passengers to their originally intended destination. Upon landing at the original destination, the dispatcher would provide or coordinate any required feedback necessary for the airline's quality control processes and continuous improvement.

Today, dispatchers collaborate with and receive support from other specialists (i.e., other actors in the AOC). These include:

- Operations Managers: These senior level dispatchers provide support around flight cancellation and delay decisions.
- Traffic Management: This unit acts as the point of contact for the ATC system.
- Crew Schedulers: In the event of (for example) a flight delay or cancellation, crew schedulers manage the scheduling of pilots and flight attendants. They also oversee training, vacations, monthly scheduling and monitoring adherence to on-duty time restrictions.
- Maintenance: As scheduled or un-scheduled maintenance issues arise, this group works with dispatchers to coordinate resources (e.g., supply chain and manpower) to address them.
- Meteorologists: Given that weather is a critical factor in flight planning and operation, dispatchers work closely with the meteorology department to gain insight on weather forecasts and en-route conditions.
- Medical Personnel: The medical personnel team assists dispatchers to coordinate medical support during in-flight incidents or emergencies.

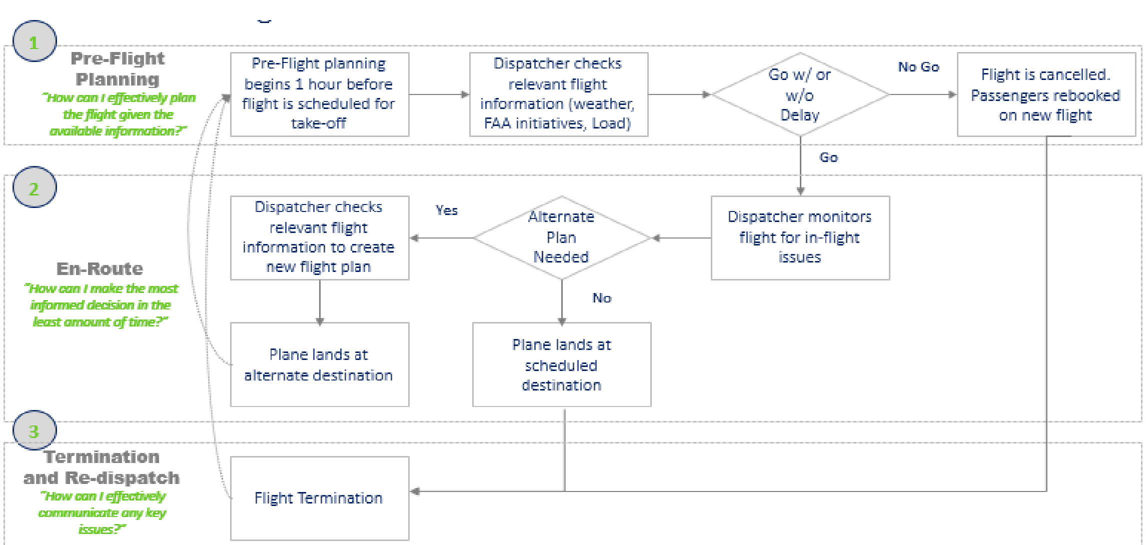

Figure 2: Dispatcher Operations

A dispatcher also uses a constellation of information technology systems, and a variety of information sources including a digitized reference library containing documents of a wide variety, with information that might be needed during one or more phases of a flight.

This collection includes, but is not limited to:

Table 1: Data Sources

Operations Data and Manuals	Contains procedural information for airlines.
Aircraft Manuals	Provides operating limitations, procedures, and systems information for flight crews for each airframe in the fleet.
Quick Reference Handbook (QRH)	Contains checklists for events a flight could experience (abnormal events, engine fires, Engine-indicating and Crew-alerting System, etc.). Also called Emergency and Abnormal Checklist (EAC).
Master Minimum Equipment List (MMEL) for all airframes in fleet	Lists all equipment on an aircraft type and details which equipment is allowed to be inoperative without grounding the aircraft.
Notices to Airmen (NOTAM)	Contains streaming, expiring information concerning the establishment, conditions, service, hazard, events, etc. (runway shut down, flights of important people, military restrictions, etc.).
Meteorological Terminal Air Reports (METARs)	Contains weather forecast that include: the airport identifier, time of observation, wind, visibility, runway visual range, present weather phenomena, sky conditions, temperature, dew point, and altimeter setting.

Additional sources currently in use are summarized in “Appendix B: AOC Data Sources” on Page 20.

Spanning all phases of a flight, dispatchers and other AOC decision makers navigate a work flow, containing complexities and variations below its top level. At that level, the flow consists of three events or states: stimulus, decision, and action. Stimuli come in a variety of forms ranging from simple to complex. At the simple to medium part of the spectrum are stimuli generated by individuals or systems and targeted at specific actors in the AOC. A prominent example are actor-initiated messages arriving via *data link* systems including the Aircraft Communications Addressing and Reporting System (ACARS), (from ARINCDirect—Rockwell-Collins), the Aircom system (from Société Internationale de Télécommunications Aéronautiques), or others. There are also a variety of attention-grabbing indicators in windows on console displays (e.g., something turns red or flashes). Each stimulus carries with it a context (e.g., the pertinent originator identification, nature of the problem or issue, other textual information, time stamp). These stimuli lend themselves to automated analysis to trigger information retrieval.

At the complex end of the spectrum are stimuli that are more indirect and less explicit, often taking the form of information from more than one source that, in combination, indicate a condition recognizable to a professional dispatcher (or other actor) by virtue of their expertise, experience, and intuition, which requires decision and action. These kinds of stimuli are much more varied, often more subtle, and therefore are more difficult to automatically detect.

In all cases, when a stimulus occurs, the dispatcher is faced with making a decision and choosing an action based upon the context conveyed by the stimulus, along with a much wider collection of possibly relevant information. The more supplemental information that is available immediately, the more completely informed a decision can be. But increasing the volume of available information can potentially overwhelm a dispatcher (or any human decision maker). Thus, whenever a large amount of information is deemed *potentially relevant* to a problem at hand, or when we consider increasing that collection substantially, *relevance* of information becomes increasingly important, in order to avoid overload.

Data sources come in a wide variety, and while a dispatcher has many resources to help optimize decision making, the process of using them can be time consuming and manually intensive. Data are often of limited certainty (e.g., forecasted weather does not materialize), untimely (e.g., creeping maintenance delays, poor communication with the ATC system), or too vast (numerous Notices to Airmen (NOTAMs) that need to be sorted through, and a large and expanding reference library). There are also data that the dispatchers may not even have the ability to view due to its volume or lack of accessibility.

Additionally, multiple issues can arise simultaneously, requiring dispatchers to weigh numerous options for taking action under a short time constraint (e.g., as it may relate to mechanical failures, medical emergencies, or other more urgent matters). Interrelated simultaneous decisions can conflict. Non-optimal decisions can result in inefficiencies, such as diversions to non-ideal locations that can result in higher expenses, and potentially unsafe situations.

IBM Watson can read or *ingest* an unbounded collection of natural language documents, determine relevance of every passage in every document, and respond to natural language questions posed by dispatchers and other AOC actors with a ranked list of those most relevant to their questions. The AOC, therefore, can greatly benefit from a flight operations advisor to minimize the time required to obtain relevant and complete information, enhance efficiencies, and give dispatchers and other key AOC actors more options and increased confidence in their decision making process.

The Role of Watson in an AOC

This section describes how Watson might be used as an advisor to the many actors in an AOC, and dispatchers in particular. We first describe the overall view and then discuss a multi-phase instantiation of increasing capabilities.

Objectives and Capabilities

The objective is to help AOC teams (not just individual dispatchers) respond effectively and more efficiently to air operations issues. These issues can span a wide range of domains, including some of the most frequent:

- Maintenance
- Weather
- Perturbations in ATC
- On-deck changes (both planned and emergent) at major airfields

A problem at a single major hub, for example ATL, JFK, LAX, DFW, or ORD[1], can reverberate throughout the system. Dispatchers operate in an environment where the volume of data are overwhelming, but they remain challenged to consider second and third order implications of their decisions.

Additionally, dispatchers are occasionally asked to respond to emergency conditions that represent potentially catastrophic situations when an aircraft reports issues of system malfunction or issues with passengers. In these circumstances it is imperative that AOC teams respond quickly and effectively; this means having rapid and reliable access to precisely the reference documents required to help them formulate an optimal response.

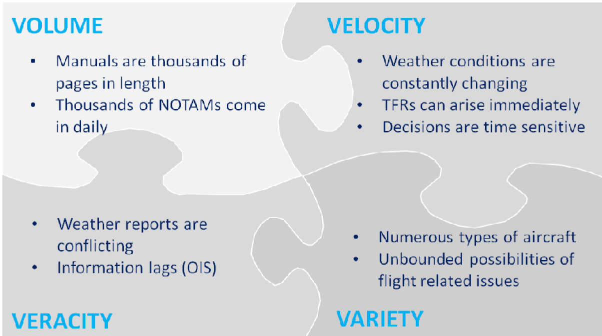

Figure 3: Large Amounts of Data in the AOC

Some problems are not strategic, such as choosing the best route given the weather or planning a fuel load, and are basic optimization problems. These predictive problems are well understood, and even though they are computationally complex, there are well-established mechanisms and tools to facilitate the right solution. They are routine, and dispatch teams solve them every day. That they rely primarily on sources of structured, even streaming data of only transient relevance (such as current runway condition or weather), means they are not necessarily a Watson problem to the degree that existing tools can handle them satisfactorily. However we anticipate these sources will be quite useful in triggering new kinds of stimuli for decision processes, setting context for questions, and other advanced uses.

---

[1] Airport codes for Atlanta, JFK International, Los Angeles International, Dallas-Fort Worth, and O'Hare airports.

More involved problems, however, present controllers with a far greater challenge.

- These problems are less well defined than a fuel or weather planning issue; dispatch teams must first understand what is happening, and then try to evaluate possible solutions from a set that is potentially unbounded.
- Reference materials span a wide range of sources, formats, and approaches to a problem. These documents are not currently aggregated in a manner to render them easily searchable, let alone under time pressure. AOC personnel most often perform manual selection of volumes and simple keyword searches of one document at a time.
- Nuggets of wisdom that might suggest a solution to an unusual problem could come from almost any source, or have meaning only when combined across documents. It is not reasonable to expect dispatch teams to have complete mastery of the quantity of data that could prove useful to them.

Beyond the intrinsic nature of these problems as outlined above, dispatchers are expected to respond on the basis of large volumes of unstructured materials contained in aircraft maintenance and operating guides, airfield procedures, federal regulations, incident reports, and a host of other documents. No human reader could master this material to the degree that would be required to derive comprehensive answers for an immediate response, or completeness in considering alternatives.

We propose a gradual integration of IBM Watson into routine AOC operations. In this vision, Watson will support reaction to emergent situations reported by, for example, data link sources such as ACARS and Delta's internal messaging service, the Operations Support System (OSS), and other explicit stimuli arising from the various sources being monitored. In principle, the objective is to improve both speed and quality of response because:

- Watson will enable expanding the size and range of materials readily at the fingertips of dispatchers and other AOC actors. At present the size is limited, and there is little federated or automated search capability spanning the collections.
- Watson will enable AOC teams to draw from a more diverse set of sources and find precisely the information that will help them resolve a specific issue.
- Watson Discovery Advisor (WDA), as the foundation Watson system upon which WFAO will be built, will expedite the location of information because it will select the most relevant passages from related documents.
- Teams will be freed from the constraint of keyword or character string matching on a per-document basis. Instead, they will be able to ask natural language questions (with context) across the entire collection and Watson's ability to extract meaning from context will enhance the impact of search.

Message Pre-processing

In the dispatcher's workflow, a particularly common type of stimulus takes the form of streaming text messages of various sorts—for example, data link messages, NOTAMs, internal text messages like Delta's OSS, and messages from ATC.

The most important of these is the data link channel. Perhaps 90% of messages from the cockpit flow via data link channels (such as ACARS) as opposed to voice channels. Many of these messages are addressed directly to the AOC, with the dispatcher responsible for a flight the primary recipient. During our AOC visits, we saw that dispatchers tend to keep at least one of their multiple physical screens tiled with ACARS message windows.

Many data link messages are aimed at entities other than the dispatcher, but the dispatcher is copied on these too and is often expected to have seen them. However, there are so many messages arriving that dispatchers often fail to review ones that they are merely copied on.

United Airlines has a system called *IONBar* that at least attempts to categorize incoming ACARS messages based on a simple keyword search. The technology is about fifteen years old and while not very sophisticated. United dispatchers nevertheless find it very useful in the same way that someone receiving hundreds of emails per day finds rules for handling inbox traffic useful for sifting the wheat from the chaff. Delta does not have a similar system; their dispatchers sift through the ACARS and other data link traffic manually.

Watson can be used to pre-process such messages in order to sort or prioritize them, or to extract context and other information with which it can perform more advanced functions such as automated question composition. For more details, see ”The AOC with Watson, Phase 1c” on Page 11.

While we have focused the discussion in this section on data link messages, the same concepts apply to NOTAMs, Delta’s OSS messages, messages from ATC, etc. All could help dispatchers (and other actors in the AOC) more effectively manage the significant cognitive load in handling their incoming messages.

The AOC with Watson, Phase 1a

Phase 1a will focus on maintenance-related issues, as these are among the most frequently occurring. Watson will allow AOC actors to optimize decision making when mechanical issues arise, by aggregating information sources, providing them with relevant questions and answers, ranked list of results, and alerts.

Watson will stand separate from other systems but nearby in a console environment. ACARS (and other data link) messages will continue to arrive as they normally do. When a problem is received, dispatch center teams will access Watson from their consoles, compose a question, and interpret the message in light of the information gleaned from Watson’s response, composed from all the documents in its corpus.

Dispatch teams receive a wide range of inputs via the data link messaging systems. Many of these are for unusual circumstances in the aircraft to which teams must react quickly. We propose using Watson to navigate relevant reference materials to present key passages to dispatch teams alongside each original message, and links to full documents if needed.

We suggest that the AOC use case presents an opportunity to begin NASA's examination of the use of Watson in aeronautics using a very representative decision process. The Flight Operations Advisor is an appropriate and technically achievable solution to a set of concrete, very real problems in AOC management. We envision using WFAO to help dispatch teams formulate responses to issues received via data link messaging systems. ACARS, for example, includes automatic reporting by aircraft systems, but also offers the flight crew the opportunity to relay free text to ground operations. It is a mechanism by which ground staff can remain abreast of developments aboard aircraft in flight.

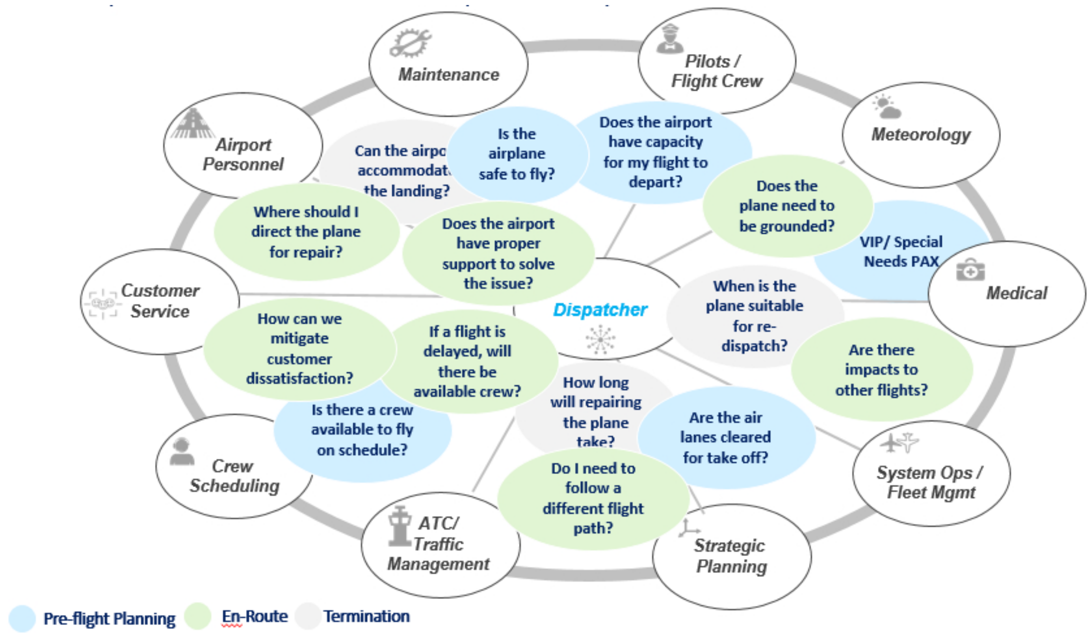

Figure 4: Many Decisions are Impacted by Maintenance Issues

Responding to data link messages is a challenge; the range of issues they can address is, in principle, unbounded. Information relevant to these messages could come from almost anywhere, and by definition, the most pressing problems or emergencies are likely to be the most challenging, because ground crews are less likely to be experienced in responding.

We expect to address this challenge in a series of steps, which will gradually increase in complexity and sophistication as both our delivery team and dispatch crews become more sophisticated and experienced with the solution. The phased solution (see below) offers significant advantages:

- It leverages the ability of WDA, and WFAO built upon it, to serve as a highly knowledgeable research assistant. Watson will effectively memorize a huge volume of documents of direct and peripheral significance to the management problems faced by AOCs. It will afford dispatchers and other actors the ability to make well-supported decisions in reaction to hard problems in less time, with greater confidence.

- It will provide the AOC teams with the most relevant excerpts from the most relevant documents, accompanied by links to the full source documents.
- It will expand the range and volume of documents being searched in an emergency, and allow decision makers to consult expertise from related fields where a similar problem might provide a key insight for a decision. It can identify related information because the ability of Watson to extract deep meaning from context frees it from reliance on keywords.
- Our initial proposal involves manually written Watson queries based on information contained in ACARS messages from aircraft. However, in subsequent phases we can envision a larger system in which key ideas extracted from data link messages can automatically inform the query process. The dispatcher would automatically receive relevant initial responses, with citations of appropriate authority, simultaneously with the message—and still have the ability to follow-up with more detailed or subsidiary questions.

The benefits are comprehensive. Watson will:

Create Time Efficiencies- By presenting information in a consolidated view, Watson will enable dispatchers to make decisions quicker, translating into less delays, cancellations, and impact on flight crews.

Generate Cost Savings- By helping generate insights, Watson will enable dispatchers to optimize route planning, thereby effectively managing fuel use and other operating costs.

Increase Safety- With easily searchable content, Watson will allow dispatchers to make more accurate, timely decisions, meaning avoiding potentially catastrophic events and improving safety.

Increase Customer Satisfaction- Because Watson will enable quicker and more accurate decision making for dispatchers, it can help improve proactive flight disruption planning, which will result in better customer satisfaction**.**

The AOC with Watson, Phase 1b

In Phase 1b, we envision an integration of ACARS (as an example of a data link channel) and Watson *on-the-glass*; that is, the messages will display on the same screen within which the dispatcher can also access Watson. Although the systems will remain functionally separate, and the entire operation will be manual (in the sense that the dispatcher still reads messages and formulates questions), this relieves the controller of the burden of switching attention between systems—a useful attribute in situations where response time is critical—and is the first step towards integration with existing systems.

In much the same way we will increase the complexity of the integration between data link channels (beginning with ACARS) and WFAO, we suggest that we should approach ingestion with an eye towards beginning with those documents that will ingest easily, and gradually increase the complexity of the process as the program matures.

Phase 1b will continue to focus on the dispatcher and his handling of maintenance issues.

The AOC with Watson, Phase 1c

We believe that the next logical step in the integration process is automatic extraction of key ideas from data link messages, enabling automated query generation for WFAO.

In Phase 1c we enable automated query generation. Watson can screen data link messages for key words or phrases and interpret their meanings. Each can then be used to generate a question, selected from a defined list of frequent questions and tailored such that when the controller first reads the message she also has immediate access to the passages returned by the Watson query. Though this may be irrelevant for many searches, it will be a huge time saver in cases where circumstances demand immediate action. AOC teams will have a summary of the most relevant material to formulate a response, along with links to original documents from which the passages were taken.

The actor in focus during Phase 1c will continue to be the dispatcher, but we will expand the scope to include a selection of issues beyond maintenance, and therefore also expand the subset of actors in the AOC with whom the dispatcher interacts. Candidates for this expanded scope include those interactions, which occur most frequently, for example weather, airports, and ATC.

It seems likely that weather-related issues can be addressed in Phase 1 considering the recent IBM acquisition of The Weather Company, which will enable easy access to their data and services. This may take the form of displaying current relevant weather conditions associated with a flight in question (i.e., no integration with Watson per se) within the same user interface as the Watson system.

Added Function

Streams

We envision employing streams to provide relevant contextual information to controllers to help in their mission to formulate a response to aircrew. See “Added Function” on Page 13 for details.

Watson Explorer

Data link messages tend to be terse and use lots of acronyms and specialized jargon, and they are often not structured in complete sentences and paragraphs. They are quite similar to instant messages in this regard. IBM has had good success using Watson Explorer (WEX) Content Analytics to pick apart instant messages, tweets, and other social media chatter, so IBM has a basis for confidence that WEX can make sense of data link message traffic. WEX can handle data streaming in incrementally. It can identify message topics. It can be loaded with a dictionary of synonyms so that topics described in different language can be correctly resolved. WEX can use a topic hierarchy so that, for example, references to “rain” and “snow” roll up under “weather,” references to “engines” roll up under “maintenance,” and so forth.

Thus, WEX can sort the incoming data link messages by topic and, to some extent, categorize them according to qualities such as urgency, severity, or seriousness. This alone has value, as evidenced by United’s *IONBar* system. Beyond this, in later phases of the project, WEX could be used to automatically call up relevant manual sections in response to topics raised in the message traffic. The reference in the data link message could be converted to a link so that dispatcher need merely click on it to bring up the relevant manual sections or passages. Such pre-processing of messages could also automatically trigger a common or frequent question to Watson and display the top ranked responses.

Watson therefore has the ability to sort and prioritize the incoming message traffic, which would eliminate much of the manual sorting and prioritizing that currently occurs. And Watson has the ability to preposition the relevant sections from the maintenance manuals, flight operations manuals, NOTAMs, and so forth, thus eliminating the manual keyword searching of individual manuals that occurs today.

The AOC with Watson, Phase 2

Phase 2 will expand the function of the Flight Operations Advisor in three ways: (1) inclusion of structured data into Watson’s knowledge base and question processing, (2) further expansion of the corpus of documents in the knowledge base, and (3) support for actors in the AOC beyond the dispatcher, the main focus of Phase 1 (a, b, and c), ultimately evolving into a decision support system spanning the AOC and its many intertwined actors and processes.

Structured Data

Structured data relevant to AOC decision processes falls into two categories: data at rest, meaning data that exists in relational databases and other similar collections; and streaming data, meaning data that is provided as a continuous stream of updates, persistent (i.e., stored) or not, ranging widely in volume and frequency.

Structured data at rest can be a source of fact-oriented answers to questions, and can serve to provide context peripheral to a question. Examples are database tables containing historical records of flights, data about fleets and crews, and weather information.

Structured streaming data can also be useful as a source of answers, when a cumulative state is kept and updated in real-time in anticipation of its use in answering questions, help in establishing context, and as input to separate evaluation for patterns that might trigger alerts or other complex stimuli (referred to in “The AOC Today” section as indirect or non-explicit stimuli, usually triggered or initiated by a human expert). Examples include real-time, in-flight aircraft data, airport updates, and weather alerts.

Corpus Expansion

Expansion of Watson's corpus or knowledge base is a continuous act of data source curation and ingestion. The objective is to consider a wide variety of information sources beyond those presently in use by AOCs, but possibly relevant to a variety of decision-making processes. An initial selection of sources to be considered appears in "Appendix C: Additional Possible Sources" on Page 23 and includes significant data sources from government, industry, and aeronautics trade media.

Spanning the AOC

The major, long-term objective of Phase 2 is to evolve WFAO to support the AOC as a whole, which is seen as a multi-actor, multi-process, collaborative, decision making entity that also interacts with outside entities.

Added Function

To this point the focus of this document has been on using Watson to discover the most relevant information (to a given issue) across an expanding collection of sources. Appropriately, we have been considering mechanisms by which NASA and airline teams can enhance service delivery and improve responsiveness in emergent situations through application of WFAO and WDA services to a corpus of information vastly larger than could be accessed by a human analyst. WFAO will significantly augment the decision making capacities of expert, experienced humans in the loop by relieving them of the need to manually seek out the material best suited to address the need at the time it arises.

To support the expansion of capability in Phase 2 as described above, additional function must be added to the architecture. While we have not yet begun a detailed design for Phase 2, it is clear that particular functions will be useful in building these new capabilities. For the purposes of this paper, we focus on two: the Knowledge Graph (KG), a function within the Watson system today, and Streams, a platform for analytics external to Watson to pre-process high velocity data streams. Each is described briefly in the following sections.

Knowledge Graph

The latest capability for WDA is the KG. It affords a capability for the decision maker to extract information about the relationships between entities. It is a powerful tool of discovery in that it allows the user to expose and explore connections and relationships she might not have known existed. A key idea is that the KG is *derived*; it melds structured and unstructured information to form a composite representation of the entities in the corpus.

The KG is a representation of known and inferred facts about the world; it is built at ingest, and contains information about entities, their attributes, and relations between entities. It is derived from multiple sources, including curated information from structured sources, and information about entities extracted from processing documents in the corpus. KG information is stored in a relational database (IBM DB2), so it has a highly scalable capacity for storing and retrieving relevant information. KG information is used in a minor way in the classic factoid pipeline; its main use is in the new *Knowledge Canvassing* functionality, a process through which an analyst may enter two or more entities and explore the connections between them.

WFAO's initial value is its ability to exploit unstructured information. However, it is clear that real world requirements drive a need for simultaneous access to related to the information derived from the unstructured text. Although it is not the only mechanism to meld structured and unstructured information, the KG is a conduit for including valuable structured information. At present only DBPedia[1] is integrated into WDA, but the capability to expand structured sources is present.

Streams

A considerable portion of the data relevant to air operations are streaming, perishable, real time information such as weather conditions, ATC congestion, runway conditions, and temporary conditions described in NOTAM bulletins. While not a candidate for ingestion into WFAO, these streams of information provide important contextual information for dispatchers and airspace managers, and they serve as valuable inputs to the kinds of optimization problems faced in the space.

Infosphere streams allow for the integrations of assorted information flows (of the type itemized above) to monitor for a linear combination of conditions that would have a bearing on choices being made by dispatchers and aircrew. We envision employing streams to provide relevant contextual information to controllers to help in their mission to formulate a response to aircrew.

Project Staging and Time Line

We propose a gradual integration of the WFAO into commercial airlines' AOCs. WFAO will support decision making in reaction to emergent situations. The objective is to improve both speed and quality of response because:

- WFAO will answer specific user questions in an unfamiliar circumstance.
- WFAO will expand the range of materials available to AOC personnel to find the information most relevant for resolving each issue. Users receive passages ranked by relevance in precise response to queries.
- WFAO returns precise information from across the entire reference library because it selects the most relevant passages from related documents.
- WFAO extracts meaning from context to answer natural language questions. Teams are freed from constraints of keywords or character string matching.

The objective of this engagement is to enable experimentation with Watson in a meaningful context enabling the complexity of the offering to increase gradually in

[1]DBPedia is a community effort to extract structured information from Wikipedia and make it available on the Web. See wiki.dbpedia.org.

several stages. We plan a steady progression from a simple, stand-alone instance of WFAO to a system where WFAO functionality is integrated with other systems to afford dispatchers the most complete possible resource to manage the national airspace. This plan is detailed in Table 2 below, although it is important to note that specific plans are expected to evolve to reflect the experience gained in each preceding iteration.

System training takes place in each phase, and takes the form of question-answer pairs written by subject matter experts (e.g., from NASA, consultants, or other aviation industry actors). This process is not intended for Watson to learn answers. Rather, when trained on such material, Watson's machine learning models learn *how* to answer questions more precisely, judge confidence, and score and rank evidence for relevance more accurately. A new round of training takes place whenever significant changes are made to the corpus or functions are added that affect how questions are interpreted or documents are annotated (i.e., how Watson's knowledge base is built).

Phase 1

Phase 1 will be our first kickoff to having NASA explore WFAO (and the underlying WDA capabilities) with relevant ingested content such as MMELs and aircraft maintenance manuals. This phase will be a limited user trial, deployment of Watson Discovery Advisor on an IBM hosted platform, non-production environment. Phase 1 will be delivered in three phases tentatively labeled 1a, 1b, and 1c, where each phase increases the capability of the system, and expands the diversity of ingested data.

Phase 1a is the initial phase of Watson engagement at Ames. Corpus selection will be guided by simplicity of the ingestion process, and the WFAO interface will be completely separate from other available systems (including ACARS as an example of a data link message channel). This phase is intended to offer a preliminary look at the capabilities of WFAO in an environment where there is currently limited search capability over very technical information, and it is intended to influence development of subsequent phases.

Phase 1b will slightly increase the complexity of 1a by offering integration of ACARS and Watson "on the glass". Although the WFAO and ACARS systems are functionally distinct, this phase is intended to facilitate joint usage, and expand the audience for the service.

Phase 1c expands the overall architecture to include WEX for the purpose of preprocessing the data link messages. We envision a system where keywords are extracted to enable automatic generation of WFAO queries. This capability is intended to provide the dispatcher with the data link request for assistance with a set of passages ranked by relevance to the problem; in effect, the dispatcher will simultaneously receive both the question and a set of possible answers. The depth and breadth of the ingested corpus will continue to increase.

Finally, as part of Phase 1, the team will plan and document Phase 2.

Phase 2

Phase 2 is still generally undefined, because it is expected that requirements will evolve throughout the stages of Phase 1. We can broadly identify the objectives of this phase, all of which are designed to increase the integration of WFAO with other systems, and enhance dispatcher situational awareness over the national airspace.

- Move system from a test environment to a production environment.
- Integrate Watson with streaming, perishable information such as weather conditions, runway visibility, airport traffic congestion, and temporary conditions at relevant airports.
- Integrate Watson with more permanent structured information (such as aircraft performance charts or airport parameters) through inclusion in the KG.

Table 2: Project Phases

	Decision Optimization 1A	Decision Optimization 1B	Decision Optimization 1C	Decision Optimization 2
Phases	Phase 1A – Dispatcher and Maintenance	Phase 1B: ACARS/WDA *on the glass*	Phase 1C : Automatic query generation	Phase 2: Integration
Description:	Watson will enable dispatchers in maintenance related decision making, by aggregating data and providing Q&A responses related to ACARS messages.	ACARS messages are displayed together with WDA responses.	Watson will extract key terms from ACARS messages to generate automatic queries. Expand topics beyond maintenance.	Watson integrates streaming and structured data with unstructured text
Interactions	Primary Actors- Maintenance Secondary Actors- ATC, Airports, Crew Management, Flight Crews	Primary Actors- Maintenance Secondary Actors- ATC, Airports, Crew Management, Flight Crews	Primary Actors- Maintenance Secondary Actors- ATC, Airports, Crew Management, Flight Crews	Primary Actors- Maintenance Secondary Actors- ATC, Airports, Crew Management, Flight Crews
Functional capabilities	• Dispatcher interacts with maintenance & other peripheral actors • ACARS acts as primary stimulus • Concentrated list of questions • Links and passages will be returned	• Dispatcher manually reviews ACARS • WDA accessible with ACARS messages	• WEX interprets ACARS messages. • Automatic queries generated, messages displayed with relevant passages	Streaming and structured data combined with WDA output to enhance decision making
Solution Components	• WDA	• WDA	• WDA • WEX	• WDA • WEX
Key Data Sources	• ACARS • MMELs • Aircraft Manuals	• ACARS • MMELs • Aircraft Manuals	• ACARS and other messages • Reference Library subset	• ACARS and other messages • Reference Library subset

Appendix A: Links to Other NASA Watson Projects

As background, in this section we summarize some other Watson uses being considered at NASA. There will undoubtedly be some overlap of design, corpus and function to consider as each one is developed. That should enable each to learn and benefit from the experience of the others. There may also be an opportunity to share an implementation platform to reduce costs.

NASA Armstrong – Flight Advisor

In this use case, WDA provides flight advisor services, in many ways similar to the Watson Flight Operations Advisor. But instead of a decision support for ground operations, it supports flight crews, particularly pilots. Watson would provide the core of a broader flight advisory system, conceived as a mechanism to help reduce pilot confusion in dynamic and unanticipated situations, of varying degrees of urgency.

WDA can perform the role of trusted assistant to help pilots discover information most relevant to a currently unfolding scenario among a potentially huge set of natural language documents (a corpus) read into its knowledge base (curated and ingested) from existing data sources.

Example data sources:

- Aircraft and specific aircraft system manuals
- Technical bulletins for aircraft/systems
- Incident/accident reports
- Industry journal articles on flight safety
- Government directives and reports on aviation safety
- Specific flight data recorder information for aviation mishaps

WDA can form the basis of a much larger ecosystem, where its ability to search and discover information in a large body of unstructured text can be extremely useful in conjunction with other systems.

WDA can search a corpus to help an analyst determine a set of probable outcomes under a set of circumstances. This ability has been successfully employed to calculate prior probabilities for Bayesian predictive models, and to influence the distribution of outcomes at nodes in simulation models.

Watson can be used in conjunction with models for analysis of streaming data in near real time so that pilots can be provided with corrective advice derived from manuals and past incidents when analysis of real time aircraft systems data suggests a potential problem.

WDA can be part of a deep learning ecosystem to augment the response of autonomous flight systems to emergent, unanticipated scenarios.

NASA Langley – Four Use Cases

Langley 1: Aerospace Innovation Advisor

In this use case, WDA supports discovery of technical information for writing or reviewing proposals, reports, and other publications and project-related documents. The system would potentially ingest the entirety of NASA's own technical publications, plus any external licensed sources (professional journals, technical publications, news, etc.). This use case can also support the engineering design process. It could conceivably provide additional flight operations information, too.

NASA has researchers engaged in efforts to identify transformational breakthroughs. They seek technology that will have a major impact.

Researchers tend to have an expertise that is narrow but very deep. It is difficult for them to exploit their expertise in a broad and general way across many different fields. This use case seeks to use Watson as a tool to help researchers generalize their results. Watson would help researchers identify similar work to their own in related disciplines, and see how it might be leveraged—allow them to look for research in other fields that is similar to their own, and follow this information to exploit breakthroughs in other fields that may well be relevant in their own work.

Researchers would be on the lookout for people and organizations that are doing something relevant to their own research. The idea is that there is good information in peripheral areas, where the researcher is less likely to know them or look for information. This will allow them to dive deep into that research. Additionally, Watson is seen as a vehicle to help rank-order research by strategic significance. The proponent of this use case suggests that Watson could be used to identify strategy documents that set priorities that could be used to select between competing projects.

Langley 2: Discovering What Is Missing

Researchers are very good at following leads, if they can find the leads. Instead of using simple tools like Google to identify leads, Watson could perform the task of monitoring new publications for specific new work pertinent to a given field of study or project.

In this role Watson would help make connections across sources, not confined to any one specific field. For example, in the earliest stages of a project, it would provide an ability to search massive numbers of abstracts to identify potentially valuable avenues of inquiry.

This is strongly tied to the Langley 1 use case.

Langley 3: WDA for Human Health Effects

In this use case, Watson is an advisor for investigating effects and mitigation of phenomena related to space environment radiation health effects.

Since there are relatively few astronauts, this is an area where there is limited access to data. A primary researcher indicated that she has about 1,000 documents for an initial corpus. Watson would exploit related literatures, such as those that explore connections between radiation exposure and incidence of cancer in the terrestrial population, and possible similar connections in the pharmacology, and WMD scholarly literatures. NASA's interest is in identifying countermeasures to radiation in order to prevent radiation caused cancer rather than cure it. Countermeasures are different from treatment strategies, but there is likely useful information in the medical literature on cancer treatment that might help research efforts.

In effect, the intent is to explore everything that is known about radiation in the terrestrial literature, and map it into research on space exposure.

Langley 4: Remote Medical Diagnosis Advisor

Remote diagnosis for long duration missions (e.g., International Space Station assignment).

Appendix B: AOC Data Sources

Name	Abbr.	Description	Access	Length	Format	Updated	Structure	Source Owner	Specific	Source
Airline Operations Specification	N/A	Contains airline operations information	Pull-Manual	4 chapters, 100-150 pages	PDF	TBD	Paragraphs	Airline proprietary	To Airline	Airline proprietary
Operations Data Manual	N/A	Contains procedural information for airlines	Pull-Manual	Several hundred pages	PDF	TBD	Paragraphs	Airline proprietary	To Airline	Airline proprietary
Flight Operations Manual	N/A	Procedures that pilots must follow (checklists, etc.) specific to an airline	Pull-Manual	Varies	PDF	TBD	Paragraphs	Airline proprietary	To Airline	Airline proprietary
Aircraft Manuals	N/A	Provides operating limitations, procedures, and systems information for flight crews to operate a specific plane	Pull-Manual	1500 pages	PDF	TBD	Paragraphs, diagrams, charts	IBM can access	To airplane	https://www.redskyventures.org/doc/other-poh/Boeing_737-600-700-800-900_Operations_Manual.pdf
Quick Reference Handbook	QRH	Contains checklists for events a plane could experience (abnormal events, engine fires, EICAS, etc.). Also called EAC	Pull-Manual	TBD	PDF	TBD	Checklists	NASA to provide (link does not work)	To airplane TBC	http://www.flightdeck737.be/wp-content/uploads/2010/03/B738W-QRH.pdf
Airport Winter Ops Plan	N/A	Provides guidance for airport staff on how to navigate through winter weather (how to remove snow, where to move snow, how to notify air carriers of snow)	Pull-Manual	100 pages	PDF	TBD	Sentences and diagrams	IBM can access	To airport	http://www5.passur.com/mke_docs/mkepdf3.pdf
Federal Aviation Regulations – Aeronautical Information Manual	FAR AIM	Contains the fundamentals required in order to fly legally in the country of origin. It also contains items of interest to pilots concerning health and medical facts, factors affecting flight safety, a pilot/controller glossary of terms used in the ATC system, and information on safety, accident, and hazard reporting.	Pull-Manual	~1000 pages	Web browser (PDF possible but always being updated)	Constantly	Paragraphs, charts, diagrams	IBM can access	To USA	http://www.ecfr.gov/cgi-bin/text-idx?c=ecfr&tpl=/ecfrbrowse/Title14/14tab_02.tpl
Master Minimum Equipment List for all airframes in fleet	MMEL	Lists all equipment on an aircraft type and details which equipment is allowed to be inoperative without grounding the aircraft.	Pull-Manual	~250 pages	PDF	TBD	Specific table/ chart style	IBM can access	To airplane	http://fsims.faa.gov/wdocs/mmel/b-737_rev%2057.pdf

Name	Abbr.	Description	Access	Length	Format	Updated	Structure	Source Owner	Specific	Source
Notices to Airmen	NOTAMs	Contains streaming, expiring information concerning the establishment, conditions, service, hazard, events, etc. (runway shut down, flights of important people, military restrictions, etc.)	Web browser	"Tweets"	Web browser search engine	Constantly, info expires	"Tweets"- lines of text. Contain specific abbreviations (ex: RWY= runway)	IBM can access	To airport	https://pilotweb.nas.faa.gov/PilotWeb/
Temporary Flight Restrictions	TFRs	Alert of short term airspace closures or limitations (President, VP flying), or certain special events	Web browser	"Tweets"	Web browser	Constantly, info expires	Sorted by date, NOTAM, Facility, State, Type, Description	IBM can access	To airport	http://tfr.faa.gov/tfr2/list.html
Meteorological Terminal Air Reports	METARs	Contains weather forecast that include: the airport identifier, time of observation, wind, visibility, runway visual range, present weather phenomena, sky conditions, temperature, dew point, and altimeter setting.	Web browser	"Tweets"	Web browser	Hourly	Coded in abbreviations (FAA standard formatting)	IBM can access	To airports at a specific time	http://www.aviationweather.gov/metar
Terminal Aerodrome Forecasts	TAFs	Contains future airport forecasts over next 18-24 hours	Web browser	"Tweets"	Web browser	Every 6 hours	Coded in abbreviations (FAA standard formatting)	IBM can access	To airports at a specific time	http://www.aviationweather.gov/taf
Pilot Reports	PIREPs	Contains pilots reports on weather data while in flight	Pull	"Tweets"	Web browser	TBD	Coded in abbreviations (FAA standard formatting): At a minimum the PIREP must contain a header, aircraft location, time, flight level, aircraft type and one other field.	IBM can access	To location	http://www.aviationweather.gov/airep
Significant Meteorological Information	SIGMET	Contains a weather advisory that has meteorological information concerning the safety of all aircrafts (e.g., significant turbulence, thunderstorm)	Pull	N/A	Web browser	As needed	TBD	TBD	To location	http://www.aviationweather.gov/sigmet
Runway Visual Range	RVR	Represents the distance which a pilot can see the runway. Contained in METARs		N/A	Specific interface	TBD	Specific interface	TBD	To airport/ location	http://rvr.fly.faa.gov/rvr/help.html

Name	Abbr.	Description	Access	Length	Format	Updated	Structure	Source Owner	Specific	Source
FAA Operational Information System	OIS	Real-time airport delay information as received from FAA facilities. The OIS system is a web-based application that displays up to the minute Ground Delay, Ground Stop, Deicing, and general airport delay information.	Web browser	N/A	Web browser	5 minutes	Dashboard	IBM can access		http://www.fly.faa.gov/ois/
FAA Aviation Information System	AIS	Information on operation status of flights	Web browser	N/A	Web browser	Real time	Various	IBM can access		https://www.fly.faa.gov/ais/jsp/ais.jsp
Air Traffic Control System Command Center Advisories	ATCSCC	Real-time advisory information as received from FAA facilities.	Web browser	Varies	Web browser	Real time, expires after 15 days	Tweet with abbreviations	IBM can access	To USA/ Canada	http://www.fly.faa.gov/adv/advAdvisoryForm.jsp
ATCSCC Reroutes	N/A	Displays all current (active) reroutes, with a link to the associated advisory.	Web browser	Varies	Web browser	As needed	Coded in abbreviations (FAA standard formatting)	IBM can access	To flight	http://www.fly.faa.gov/ratreader/jsp/index.jsp
Airport Arrival Demand Chart	AADC	Real-time airport arrival demand information as received from FAA facilities. The AADC chart displays the latest arrival demand metrics for selected airports.		N/A	Chart	Constantly	Chart	IBM can access	To airport	http://www.fly.faa.gov/Products/AADC/aadc.html
Aviation Advisory Circulars	AC	Provides guidance on compliance for airworthiness regulations		Several hundred pages	PDF	TBD	Paragraphs	IBM can access		https://www.faa.gov/regulations_policies/advisory_circulars/
Aviation Safety Reporting System	ASRS	Voluntary database that contains safety information/ incidents provided by pilots and other aviation frontline personnel	Pull	Varies by report	Web browser	TBD	Paragraphs	IBM can access	To incident	https://titan-server.arc.nasa.gov/ASRSPublicQueryWizard/QueryWizard_Filter.aspx

Appendix C: Additional Possible Sources

Name	Description	Format	Source
Government			
FAA Accident & Incident Data	Structured and unstructured data about accidents and incidents	HTML, MP3, REST API	http://www.faa.gov/data_research/accident_incident/
FAA Aviation Data & Statistics	Structured and unstructured data utilization, on-time stats & delay causes, planning data, airports, civil airmen	HTML, MS Word REST API	http://www.faa.gov/data_research/aviation_data_statistics/
FAA Aviation Forecasts	Aerospace and Terminal Area Forecasts	HTML, PDF	http://www.faa.gov/data_research/aviation/
FAA Passengers & Cargo	Stats on passenger boarding, Passenger Facility Change (PFC) reports, unruly passenger stats		
FAA Safety	Accident and Incident Data, Aviation Safety Information Analysis and Sharing, FAA Quarterly Enforcement Report, Lessons Learned from Transport Airplane Accidents, National Wildlife Aircraft Strike Database, NASA/FAA Flight Test Safety Database, Safety Data Library, Safety Record of Airlines and Aircraft	HTML, PDF	http://www.faa.gov/data_research/safety/
FAA Data & Research	Misc. publications pertaining to airports, air traffic, commercial space transportation, environment, modernization	HTML, PDF	http://www.faa.gov/data_research/research/
FAA Regulations & Policies	Advisories & guidance, regulations (FARs), handbooks & manuals, orders & notices, policy and guidance, NOTAMS, Temporary Flight Restrictions (TFR), ADs, ACs	HTML, PDF	http://www.faa.gov/regulations_policies/
FAA Training & Testing	FAA-industry training standards, pilot training, mechanic training, aircraft dispatcher training, Flight Standards district offices	HTML, PDF	http://www.faa.gov/training_testing/
FAA Air Traffic	Aeronautical Information Manual, International Flight Information Manual, Aeronautical Information Publication, NOTAMs (Class II), Orders and Notices: JO 7110.65, air traffic control, location identifiers, contractions	HTML, PDF	http://www.faa.gov/air_traffic/
FAA Airports	Standard operating procedures, statutes & regs., compliance, safety & cert., obstructions, environmental issues, data & stats., planning & capacity, regional guidance, runway safety	HTML, PDF	http://www.faa.gov/airports/

Name	Description	Format	Source
FAA Aircraft	Registration numbers, airworthiness certs., advisories, MMEL, Parts Manufacturer Approval , Technical Standard Orders, Type Certificate Data Sheets, Suspected Unapproved Parts, Repair Stations, Recreational Aircraft, Advisory Circulars, Air Transportation Oversight System	HTML, PDF	http://www.faa.gov/aircraft/
NTSB Accident Reports	All reports in various sort order	PDF	http://ntsb.gov/investigations/AccidentReports/Pages/aviation.aspx
NTSB Safety Studies	Studies, alerts, recommendations	PDF	http://ntsb.gov/safety/safety-studies/Pages/SafetyStudies.aspx
NTSB Aviation Investigations	All in various sort order		***http://ntsb.gov/investigations/Aviation/Pages/default.aspx***
Canadian Aviation Regulations	CARs, searchable		https://www.tc.gc.ca/eng/acts-regulations/regulations-sor96-433.htm
European Aviation Safety Agency	Regs. and other documents		https://easa.europa.eu/document-library/regulations

Aviation Industry News

Name	Description	Format	Source
Aviation Week	Commercial, defense, and business aviation news.	HTML, RSS	http://aviationweek.com/
FlightGlobal	"...discover everything you need to know about the global fleet of aircraft and world schedules, as well as insight on current and future market dynamics and drivers"		https://www.flightglobal.com/
Airliners.net	Portal and site. Aviation news, forums, photography: "... most visited aviation interest site on the Internet, ... users and community include airline management, frequent air travelers, aviation photographers and enthusiasts ... We strive to be the international center of aviation online."	HTML	http://www.airliners.net/
Air Transport World	"...the leading monthly magazine serving the needs of the global airline and commercial air transport manufacturing communities… Topics include airline operations, airline management issues, information technology, safety & security, eco-aviation, alliances, distribution, transport aircraft and engine programs, maintenance, repair, and overhaul (MRO), aero-politics, regulation, finance and leasing, airport development and air cargo."	HTML	http://atwonline.com/
AINOnline	Portal. Commercial, defense, aerospace, and manufacturer aviation news.	HTML, RSS	http://www.ainonline.com/
Centre for Aviation	For fee reports, visualizations, and analysis in global aviation. Sponsors events. "delivers market analysis and data that support strategic decision making at many of the world's most recognised organisations."	HTML, REST API (probably)	http://centreforaviation.com/
Aviation Today Network	Subscription based news		http://www.aviationtoday.com/

General Aviation

Name	Description	Format	Source
General Aviation News	Appears to be a "typical" news website, with aviation relevant news.	HTML	http://generalaviationnews.com/

Name	Description	Format	Source
Aero news network	Appears to be a "typical" news website, with aviation relevant news.	HTML	http://aero-news.net/index.cfm
landings.com	A catch all for other data sources. Includes links to FAA NOTAMs, Notices of Proposed Rulemaking, and emergency information. Also includes weather infomation, though this is available elsewhere	HTML	http://www.landings.com/
General Aviation News- flyermedia	Subscription to bi-weekly news letter on aviation news	HTML	http://issuu.com/flyermedia
Aviation safety and security archives	Contains a list of links of safety related aviation resources	HTML	http://archives.pr.erau.edu/resources/organizations.html#management