International Civil Aviation Organization

The Fifth Meeting of System Wide Information Management Task Force (SWIM TF/5)

Video Tele-conference, 9 – 11 August 2021.

Agenda Item 9: State, Regional and Global SWIM Updates

SEMANTICS IN THE AVIATION INDUSTRY WHITEPAPER

(Presented by Australia)

SUMMARY

This paper includes a Whitepaper discussing the role that Semantic Technologies will play in the future for the Aviation Industry and outlines a roadmap for SWIM services.

1. INTRODUCTION

1.1 The attached whitepaper reviews the current state of data standards in the aviation industry and looks at the pathway to a future of enhanced data capabilities. Specifically, the use of Semantic Web technologies and the impact these will have on defining the strategic future for aeronautical, flight, and meteorological data services.

2. DISCUSSION

2.1 The whitepaper has presented a possibility for a new holistic data-driven aviation industry underpinned by mature semantic technologies. The roadmap to this future needs wide industry consensus and planning, and promises to deliver improved opportunities for data sharing, data automation, and data knowledge for safer skies.

3. ACTION BY THE MEETING

3.1 The meeting is invited to:

a) note the information contained in this paper; and

b) discuss any relevant matter as appropriate covered by the attached whitepaper.

– – – – – – – – – – – – – –

SEMANTICS IN THE AVIATION INDUSTRY

AIRSERVICES AUSTRALIA WHITEPAPER

TABLE OF CONTENTS

Introduction	2
Current State	3
Semantics	4
Vocabulary Example	6
Model Example	8
SWIM Smarter	9
Aviation Semantic Takeoff	11
Scenario: Codeshare your Data	13
Scenario: The Departure of NOTAMs	14
Summary	16
Reference	17

METADATA

Publisher: Airservices Australia
Author: Renato Iannella
Contact: renato.iannella@airservicesaustralia.com
Status: FINAL DRAFT
Issued: 30 June 2021
ISBN: 1 74064 747 5

Photo by Ross Parmly

INTRODUCTION

The aviation industry is a global producer and consumer of data. As a safety-driven sector, the adoption of new data-enhancing technologies is always a slow and careful process, often limited by the argument that the current approach to managing data works and there is little drivers to change it. The rate at which the industry moves towards adopting new technologies at times will miss generational change in contemporary and modern advances that are commonly embraced by other industries and sectors.

This whitepaper reviews the current state of data standards in the aviation industry and looks at the pathway to a future of enhanced data capabilities. Specifically, the use of Semantic Web technologies and the impact these will have on defining the strategic future for aeronautical, flight, and meteorological data services - for both the current ATM and emerging UTM data providers and consumers

The move towards semantics provides an opportunity for the industry not only to “catch-up” to other industries, but establish a solid, reliable, and safe platform for future data services in the aviation sector.

CURRENT STATE

The current state for data standardisation in the aviation industry follows the traditional path for structured information. This is primarily an information model (expressed as UML diagrams) and a mapping to an encoding format (serialised via XML Schema). There are currently three major efforts in this area covering aeronautical (AIXM), flight (FIXM), meteorological (IWXXM) domains and others covering aerodrome mapping (AMXM). Other related industry bodies (such as IATA and the WMO) have similar efforts.

These are the "modern" efforts. The challenges for these current structured approaches for data standards include:

- The siloed approach of each domain effort reduces interoperability and reuse of common concepts.
- The UML modelling assumes a closed-world approach, that is, features and properties and tightly coupled and only have local scope and identity.
- Extensibility is difficult and requires complex tooling with little ability to verify correctness.
- Significant effort is required to maintain the UML to XML Schema mapping.
- Individual domain efforts continue to grow in scope without recognition of the capabilities of all stakeholders.

The aviation industry also has significant data defined and exchanged via outmoded mechanisms such as character-bit encoded data limited by the capability of communications technology at the time (eg ASTERIX), and messaging technologies dating back decades (eg X.400).

These are legacy systems and will most likely stay this way for a long time

The current data landscape is quite technically diverse and separated along domain functions. There is no core underlying and fundamental information model that binds the industry domains together. This represents the greatest challenge with moving forward, as without this, there are huge obstacles facing the industry to reach the goal of "global data interoperability"[1] for the safe operations of air traffic management (ATM) services. There are some mapping efforts, in the form of the ATM Information Reference Model (AIRM)[2], to provide a reference model as a first step towards shared understanding across the structured domain models.

[1] Manual on System Wide Information Management (SWIM) Concept. ICAO Doc 10039. 29015

[2] http://airm.aero

Photo by Matthew Smith

SEMANTICS

Semantics is the study of the relationship between form and meaning. For the more pragmatic and interoperability-focussed, semantics is the ability of computer systems to exchange data with unambiguous, shared meaning enabled by machine computable logic, inferencing, knowledge discovery, and data federation between information systems[3].

The Semantic Web[4] developed by the W3C is a set of technologies that provides the framework for semantic data interoperability and includes information modelling, vocabularies and ontologies, formal reasoning, and query protocols. The Semantic Web is grounded in mathematically complete logic which provides the assurance for the inferencing outcomes.

There are number of key defining characteristics to data semantics:

- Ontologies define the classes of things and properties and the relationships between them including property roles. A runway "has a" taxiway relationship and (automatically) has the inverse relationship.

[3] https://en.wikipedia.org/wiki/Semantic_interoperability

[4] https://en.wikipedia.org/wiki/Semantic_Web

- Both types of things and properties are first-class. That is, a runway and taxiway may both have a length property, but that length property is defined only once independent from both and reusable by many other things.
- Every thing has global identity and uniqueness. The length property and taxiway thing have unique identity and can be referred to and used (globally) in any other information models.
- The sameness (of things) and equivalence (of properties) can be automatically resolved. Different flight codes (eg QF1 and EK5003) can be determined to be the same (codeshare) flight.
- Ontologies are self-describing by using the sameness/equivalence feature and referring to common open ontologies. The Donlon ANSP define their Airport class as the "same as" the ICAO AerodromeHeliport class. If a system knows the latter, then they know the former too (automatically).
- Reasoning services can infer new outcomes based on the formal model underpinning ontologies. The wake turbulence category can be inferred from the weight of an aircraft without explicitly declaring this information.
- Semantic data are triples of information and commonly stored in high-performance graph databases for immense global scalability and standardised queries.

The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation
Tim Berners-Lee, Inventor of the Web

These defining characteristics are necessary requirements for semantic systems. They represent a significant uplift in information model expression and semantic automation capabilities above existing structured systems and methodologies.

VOCABULARY EXAMPLE

Vocabularies are a common way to showcase the benefits of semantics. Consider two concepts used in the aviation industry related to measuring vertical distances; "Above Ground Level" and "Mean Sea Level".

In the current state, these two measurement types are currently defined and used in the Aeronautical (AIXM) specification (see top of Figure 1). The Flight (FIXM) specification uses the code "SFC" taken from AIXM (see middle of Figure 1). The EUROCAE Minimum Operational Performance Standard for Geofencing (MOPSG) defines two attributes (see bottom of Figure 1).

Name	Data Type	Definition
SFC	string	The distance measured from the surface of the Earth (equivalent to AGL - Above Ground Level).
MSL	string	The distance measured from mean sea level (equivalent to altitude).

Attribute
Public
SFC

Details:

Notes: The distance measured from the surface of the Earth (equivalent to AGL - Above Ground Level). [AIXM 5.1]

ATTRIBUTES	ATTRIBUTES
AGL : string Public	AMSL : string Public
Height above ground/surface level	Altitude above Mean Sea Level.

Figure 1 - Current State Measurement Types

These three specifications declare these different terms (for the same concept) as "enumerated value" strings deep inside their respective XML and JSON schemas. Any consuming system would need to write code specific for each specification as there is no data interoperability, even at this structured level. Implementors could develop mappings between these codes to support local data exchange but this adds cost and effort and room for errors. This is far from ideal and so a common vocabulary would address these concerns.

In the future state, assume that ICAO is the normative organisation that defines and manages these two concepts for the aviation industry and uses a semantic approach. They would define two unique terms through Uniform Resource Identifiers (URIs), such as:

- http://vocab.icao.aero/AboveGroundLevel
- http://vocab.icao.aero/MeanSeaLevel

Typically, they would use a digital vocabulary system to manage these concepts that includes all the relevant governance and provenance metadata. This includes the formal definitions, multilingual labels, lifecycle dates, approval status, etc. In addition, the URIs that are used to provide global unique identity for the concepts, are also accessible by machine request (eg using the HTTP protocol). This allows consuming semantic systems to easily obtain the machine-readable encodings (eg in W3C RDF/OWL formats) for all the metadata describing concepts for processing and inferencing.

With these core vocabulary terms defined formally by ICAO, information standards developers can now choose to reuse these concepts directly in their specification. For example, instead of the string "SFC" in AIXM and "AGL" in MOPSG, both can now use "http://vocab.icao.aero/AboveGroundLevel" instead.

However, there is actually no mandatory requirement for the information standards developers to directly use these ICAO concept URIs in their specifications. Instead they can create and maintain their own concept URIs and either:

- Assert their concepts URIs are the "same as" the ICAO concept URIs, and/or
- ICAO assets their concept URIs are the "same as" the standards developer URIs.

Either way, consuming semantic systems can infer these relationships and use either URIs as the same concept. The feature is globally scalable. In addition, the relationship may not always be an exact "sameness" but also supports broader and narrower matches. For example, the current AIRM defines "MSL"[5] more precisely than the AIXM and MOPSG definitions. In a future state, the AIRM URI would most likely be mapped to a "narrower match" to the ICAO URI.

The "open world" nature of the Semantic Web means the definitions managed by ICAO are directly mapped/reused by any other organisations and industry sectors. The World Meteorological Organisation (WMO) has an existing semantic vocabulary system at: http://codes.wmo.int/

The WMO Codes vocabulary explicitly defines the "Mean Sea Level" concept as shown in Figure 2. The ICAO and WMO concepts are the same and can be asserted as such for semantic interoperability between the two communities.

In the wider open community, the same concept is defined by Wikipedia as: http://dbpedia.org/resource/Sea_level

This would also become a candidate for "sameness" and opens up interoperability not only globally, but across new communities of interest.

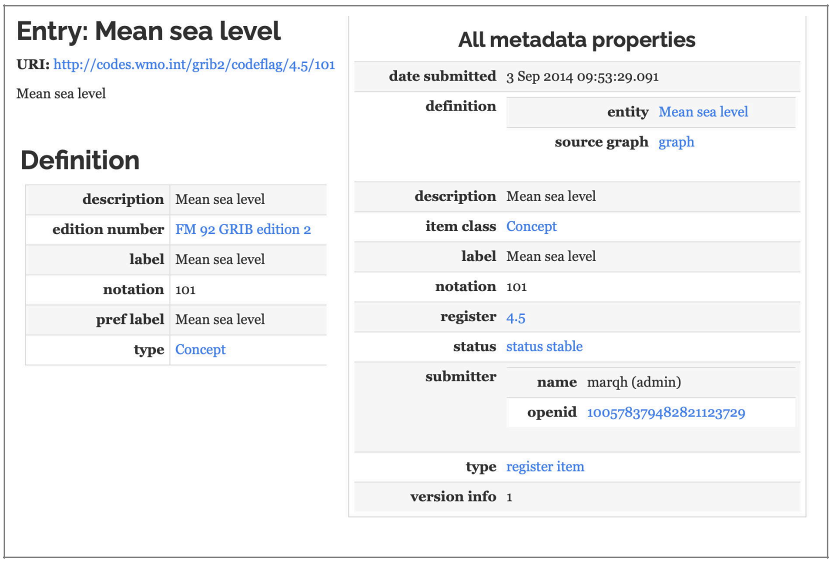

Figure 2 - WMO Codes Vocabulary Examples

[5] https://airm.aero/viewer/1.0.0/logical-model/CodeVerticalReferenceType.html

MODEL EXAMPLE

Ontologies are the key mechanisms used by the Semantic Web to define the information model and the foundation for semantic reasoning. They provide the core modelling scaffold and where more deeper semantics can also be explored and clarified for clear unambiguous meaning.

In the previous example, the EUROCAE definitions explicitly use “height” and “altitude” and AIXM uses “distance”. These may seem similar but there is always a reason why and when domain experts use these terms and these need to clearly defined in an ontology for context. Figure 3 below shows how this could be modelled in a semantic ontology.

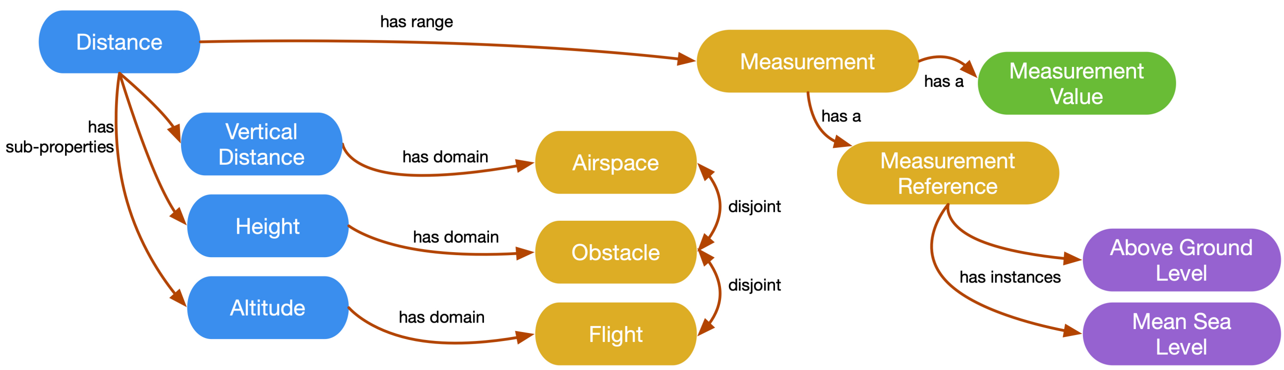

Figure 3 - Example Semantic Model

The above model includes the Distance property with three sub-properties (Vertical Distance, Height and Altitude) showing the ability to have independent hierarchically shared properties. The Distance property (and hence its children) have a Measurement, which includes the Value and two possible instances of the Measurement Reference class (Above Ground Level, and Mean Sea Level).

The three classes (Airspace, Obstacle, Flight) each are related to the specific Distance sub-property that is relevant to them (in the aviation domain). This implies that an instance of an Obstacle class should use the Height property (not the other two). Each of the classes are also disjoint, meaning the model will not allow an Obstacle to have an Altitude property.

The model also enables simple inferencing such as any thing that has a Height property can be classified as a type of Obstacle.

This is a relatively simple model but does showcase the ontology-approach to modelling. The focus is on representing the aviation domain artefacts and reusing common properties, classes, and instances, each of which has global identity and can be referred to by data specifications (in the aviation industry and beyond).

Photo by John Cameron

SWIM SMARTER

The current data interoperability approach is defined by the ICAO SWIM Concept of Operations (Doc 10039). In essence, SWIM provides broad guide-rails as to the types of modern interface protocols, data payloads, message exchange patterns, and security mechanisms that underpin the data exchange architecture. By definition, SWIM liberally allows a wide set of options to sustain the current technologies and methods used to exchange data. This does not provide a clear path to achieve 'global data interoperability' as stakeholders 'pick-n-choose' their most comfortable options.

The "System" aspect of SWIM reflects the deep history of the industry and does not necessarily reflect on the modern capability to provide real-time data. This needs to now reflect a "Service"-driven culture and, fortuitously, a potential new direction towards greater inclusion of semantic technologies in the overall SWIM environment.

The Semantic Web can provide SWIM the necessary capabilities to reach its goal of Global Data Interoperability

The current state SWIM stack is shown in Figure 4 in the "sponge" model. This highlights the difficulty posed by the numerous options at the top and bottom, with a set of constrained data payloads in the middle. SWIM stakeholders need to decide which options form their "SWIM Profile" and vendors need to support them all. This profiling concept approach forms individual and regional "siloed" conformance and poses a risk to the global interoperability goal of SWIM.

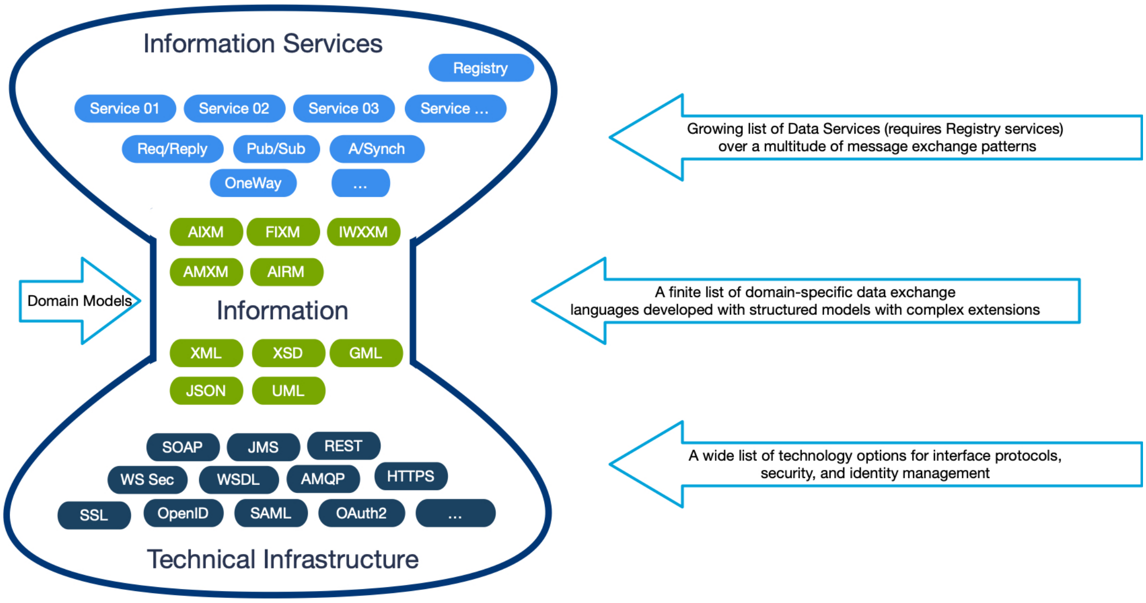

Figure 4 - SWIM Current State Framework

A future "propeller" model is shown in Figure 5. The key differences are that;

- there are fewer options as all candidates are modern well-understood technologies,
- the information services are modelled on the availability of real-time data from source systems,
- there is a core single aviation information model that provides the foundation semantics,
- there are domain models for more advanced terminology and models, and
- there is reuse of common and industry domain models to support wider semantic interoperability and access by related sectors and industries.

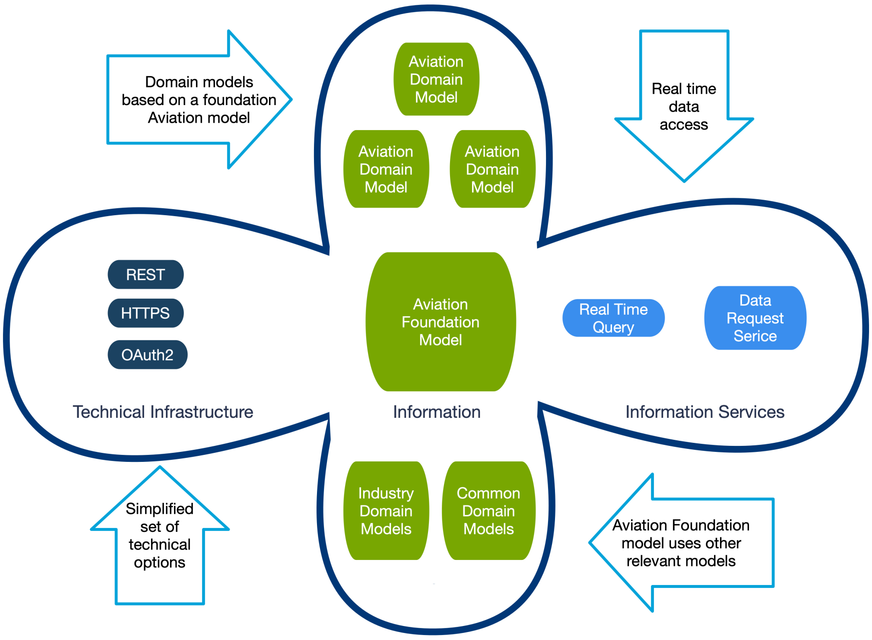

Figure 5 - Future Smart SWIM Framework

Photo by pixpoetry

AVIATION SEMANTIC TAKEOFF

Moving towards a semantic foundation for the aviation industry will provide significant benefits in operation and sharing of data for all stakeholders. The outcomes will improve understanding, insights, and development of more comprehensive aviation solutions. The industry will need a long-term roadmap for the evolution of this new capability.

Aviation is not unique in this endeavour. Finance is also global data-sharing sector and regulated industry that is maturing their semantic journey. Finance has produced the FIBO ontology[6] that has been developed by an open community process with a mission to promote a global platform-independent, machine-readable and unambiguous data standard that enables understanding of the terminology, federation and aggregation of data to improve effectiveness of decision making, to improve efficiency in regulatory conformance and to fast-track the adoption of advanced analytical capabilities.

[6] https://spec.edmcouncil.org/fibo/

The aviation industry needs to take a similar journey and develop a long term strategy and tactical plan. The roadmap will include business and technical activities that underpin the outcomes, together with a transparent and trusted governance process.

The overarching roadmap will consist of activities in the following areas;

- Communications and Outreach - includes the promotion of the new semantic technology agenda within and across the industry and dialogue with related industries bodies.
- Standards Governance Process - includes the necessary frameworks for governance approval of data standards at a global level.
- Foundation Standards - includes the development of the core Aviation Vocabulary and Ontology model.
- Domain Standards - includes the development of the advanced requirements from the industry sub-domains to support a comprehensive coverage of industry semantics.
- Regulatory Compliance - includes direct support in the Foundation models for automated compliance reporting.
- Stakeholder Collaboration - includes open mechanisms to collaborate with industry stakeholders and wider related sectors, including strong relationships with partners with mature semantic technologies.
- Systems and Services - includes all the necessary open online collaborative sharing platforms to run standards development, and the services to manage the publication of semantics vocabularies and ontologies.

The aviation industry needs leadership to take on this new challenge that will bring the aviation community together to build a solid foundation for all future data standardisation activities

The core industry stakeholders needs to take ownership of this roadmap and produce a detailed plan towards fulfilling the new semantic data future.

SCENARIO: CODESHARE YOUR DATA

The airline industry (via IATA) has began its semantic journey with the release of the first ONE Record ontology for digital cargo[7]. The vision for ONE Record is an end-to-end digital logistics and transport supply chain where data is easily and transparently exchanged in a digital ecosystem of air cargo stakeholders, communities and data platforms. In a future state, both the aviation and airline ontologies would provide a strong foundation for data sharing as well as strong data linkages.

Considering the scope of flight information from both the aviation and airline views, the linked data provided by both would provide a detailed and more comprehensive view of industry activities. Figure 6 shows and example of how the two sector views would work together. There are common elements to both sectors that are shared and well understood. In addition, each provides the additional context needed for their own specific operational requirements. All of this data is "linked" and traversable.

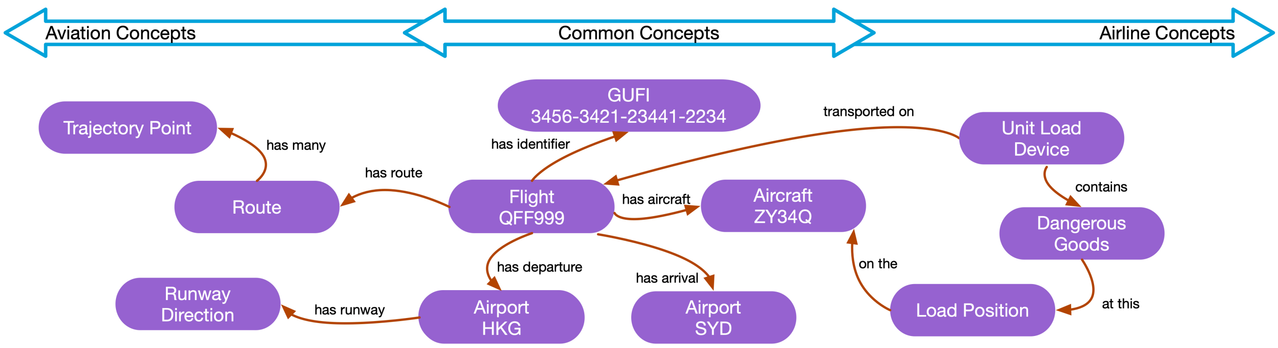

Figure 6 - Future Smart SWIM Framework

To compare to the current state, FIXM supports a rudimentary element to indicate if a flight has dangerous goods or not, which is clearly the domain of the airline as normative information. With the above approach, it is easy to now query and/or infer that flight QFF999 is such a flight, even though this was not data from the aviation sector.

[7] https://onerecord.iata.org

SCENARIO: THE DEPARTURE OF NOTAMS

The NOTAM is the aviation industries achilles' heel. It provides a necessary service to inform the industry stakeholders of all the permanent and temporary changes that are currently occurring across the aeronautical sector that may impact the safety of flights. The biggest challenges with NOTAMs is their proliferation (1.7 million active in 2020) and what this means on quickly and unambiguously understanding the relevant NOTAMs for a specific flight (a typical long range flight may have over 2,000 NOTAMs).

There are current efforts to digitise the current NOTAM, but this is not directly solving the above problems, only turning the current text-based intricate syntax into structured XML syntax based in AIXM with an event extension. There are also efforts to perform natural language processing on traditional NOTAM text.

There are longer term plans to "replace NOTAMS" and work has commenced on the Operational Concept by the ICAO AIM Working Group. The new concept will be an Operational Reporting Information Service (OPREP) that will "establish an efficient system that makes short-term changes to aeronautical information available for users" made available through SWIM services. There still seems to be the same concept in OPREPs that changes to aeronautical data need their own (separate) publication channels to operate, rather than utilising the same 'business-as-usual' services used to manage the original data.

The NOTAM system is no longer adequate to deliver the relevant short term information to the diversity of users
OPREP Concept of Operations

A fundamental aspect of the move towards SWIM is that the industry can now rely on access to real-time information with greater reliability. This will mitigate a significant number of use cases for access to temporal aeronautical data, as the past system was designed to meet the (slow) publication cycles of (paper) aeronautical products. This will mean that the consumer can simply directly ask (query) the producer for the current state of the aeronautical feature, instead of relying on old published data and searching a seperate database of NOTAMs (or even OPREPs).

Consider the scenario of a runway closure at an aerodrome for maintenance due to surface damage that will not be operational for 7 days. This new temporal status can be immediately reflected in the source dataset. This means that any real-time request for information about this aerodrome (or runway) will include the temporal change.

The main focus then of the new OPREP system then turns to how to get the most relevant data about 'things of interest' to the aviation consumer. Relevancy can be shown as the intersection of three key features; Space, Time, and Things as shown in Figure 7. Consider a flight that traverses across space and time (red thick line). Along this path, there were be a number of things of operational interest (blue dots) for this specific flight that will need to be surfaced. These things could be aerodromes, airspaces, vertical obstacles, weather conditions, etc. As well, some things would be exactly relevant (touching the red line), others may be closely relevant within a certain threshold (as shown by the pink thin lines).

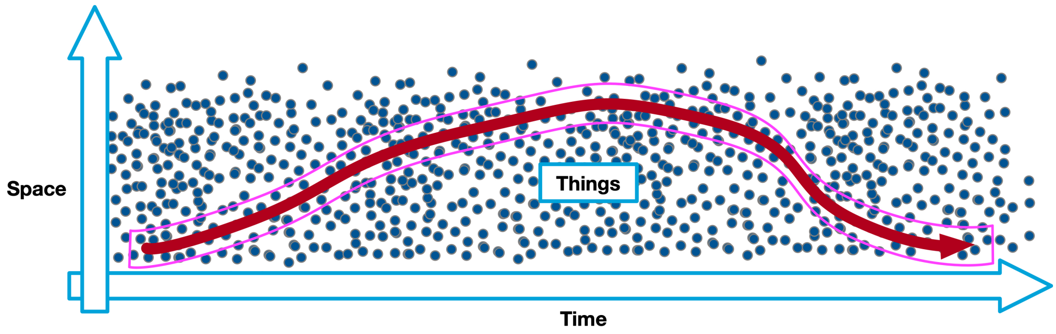

Figure 7 - Space-Time-Things Framework

Consider the current state for an airline processing NOTAMs for an international flight. Typically, each airline constantly collects all new NOTAMs published around the world (by flight regions). They then have an internal system that supports retrieving NOTAMs based on the sectors of the flight and relevant aerodromes. This is also augmented by human intervention to rank the importance of some NOTAMs over others.

The future state will be based on real-time data access with SWIM data services. Consumers can query for the state of any thing(s) at any time(s) and place(s) and be confident that the result will meet their business needs. Consider a typical Flight Plan that includes aerodromes and runways, taxi times, route (through airspace), waypoints and altitudes, weather conditions, weight, speed, fuel quantities , alternate aerodromes and routes. These are a collection of things (sometimes called features) that are of interest at a particular time. Retrieving this information in real-time will then allow for timely and transparent decision making by the airline.

In addition, with the underlying information modelled and represented as semantic data, some of the decision-making can be augmented by semantic rules and inferencing. Consider that a taxiway is closed at an aerodrome for aircraft of a specified wingspan. A semantic system can infer from the aircraft type that it has folding-wing-tips and able to use the taxiway. More complex semantic rules could be incorporated, such as determining which alternative aerodromes are valid based on the aircraft class and fuel loadings, and which alternate routes are optimal based on predicted weather observations.

The benefits of a semantics-based inferencing and rule system is the “decidability” outcomes are formally provable (as opposed to current state rule-based systems) with reliable precision. These provide the framework and confidence for an aviation future transforming from NOTAMs to real-time data services.

Photo by Nathan Hobbs

SUMMARY

The aviation industry is at the forefront of transformation and disruption. Modern semantic technologies drive transformation with new opportunities that challenge past practices with advanced information services. These technologies break-down the barriers across the aviation industry and domains by opening up data and sharing knowledge for better and safer decision management.

The industry disruption from the emerging UTM (drone) sector will highlight the technological gap between the current and future state. The use of semantic technologies will provide the common fabric for interoperability and cohesion. But this needs to be purposely managed and led in this direction otherwise a new silo is created for the industry, which it should not accept.

This white paper has presented a possibility for a new holistic data-driven aviation industry underpinned by mature semantic technologies. The roadmap to this future needs wide industry consensus and planning, and promises to deliver improved opportunities for data sharing, data automation, and data knowledge for safer skies.

REFERENCE

AIRM - ATM Information Reference Model

AIXM - Aeronautical Information Exchange Model

AMXM - Aerodrome Mapping Exchange Model

ANSP - Air Navigation Service Provider

ATM - Air Traffic Management

Donlon - A fictional aerodrome in the middle of the Atlantic ocean

EUROCAE - European aviation standards body

FIBO - Financial Industry Business Ontology

FIXM - Flight Information Exchange Model

HTTP - Hypertext Transport Protocol

IATA - International Airline Transport Association

ICAO - International Civil Aviation Organisation

IWXXM - ICAO Meteorological Information Exchange Model

JSON - Javascript Object Notation

NOTAM - Notice to Airmen

OPREP - Operational Reporting Information Service

Ontology - Formal representation of semantic models

OWL - Web Ontology Language

RDF - Resource Description Framework

Semantic Web - Collection of semantic technologies developed but the W3C

SWIM - System-Wide Information Management

W3C - The World-Wide-Web Consortium

WMO - World Meteorological Organisation

URI - Universal Resource Identifier

UTM - Unmanned Traffic Management

XML Schema - Extensible Markup Language Schema