Standards for Data Quality Assurance in ATM Modernization Initiatives

Aleksandar Balaban
m-click.aero

Charles Chen
Skymantics, LLC

ABSTRACT

Standard ATM data formats for aeronautical, weather, and flight (AIXM, WXXM, & FIXM) are becoming more widely accepted in both USA and EU ATM communities. This is evidenced in the FAA in future System Wide Information Management (SWIM) programs such as Common Support Services for Aeronautical Information Management (CSS-AIM) and Weather (CSS-Wx). In Europe, the European Aeronautical Information System (AIS) Database (EAD) and Airspace Database Repository (ADR) contain aeronautical data to be distributed via SWIM. As more and more data is being on-ramped via SWIM in the form of data publishers and web services, governance of this information becomes critical for managing the information content and quality. Especially in an operational context, the data must be validated as a precondition for safety, security, and efficiency of operations.

Since data syntax can be verified and validated through XML schema validation, the issues regarding data quality centers around data content and usage. The validation of correct data sets is essential, especially when exchanged across Flight Information Region (FIR) boundaries. A method has been developed within Eurocontrol for the checking of data content and usage using a semantic vocabulary called Semantics Business Vocabulary and Rules (SBVR), which is an OMG standard (OMG, 2015). Using SBVR, a set of business rules can be defined to enable the validation of data content and ensure data quality within the ATM data models. Furthermore, the translation of SBVR to an ISO-standard, Schematron, enables the validation through existing XSLT processes (ISO, 2006). Joint research is being conducted by FAA, Eurocontrol, and the Open Geospatial Consortium (OGC) within the OGC interoperability testbed to develop a standard process for automated data validation using current OGC standard services.

STANDARDS LOWER COSTS THROUGH RE-USE

The standardization and adoption of new ATM data models is improving the way in which producers and consumers are sharing ATM information. Using XML, the producers of data are able to exchange information using System Wide Information Management (SWIM). While SWIM is rapidly decreasing the amount of time required to connect and acquire relevant data, data quality is still a concern. Currently, it is expected that a data producer verifies and validates the information before disseminating it to the general public. However, these checks are currently limited to syntax validation or contain a limited set of business rule validations specific to the airspace in which the data is generated. This may limit the usability of the data by certain consumers, especially those external to the

specified airspace domain, even though they may be conducting operations that span across these airspaces. It can be considered to be the responsibility of SWIM to govern the validation of data for each SWIM information region as part of a larger global SWIM data exchange.

Within FAA today, producers are checked for data load, performance, and ability to maintain service level agreements. However, data verification is left to the associated program office supplying this information. For example, the FAA Federal NOTAM System and NOTAM Distribution Service (FNS-NDS) ensure that NOTAMS submitted via its service interfaces are checked via several automated business rules prior to distribution. This process of business rules enforcement is specific to the DNOTAM system, and while the architecture and process can be reused, the specific method may not be easily reusable by other data management systems. With the implementation of SWIM services such as CSS-AIM and CSS-Wx, use of a reusable validation method should be developed to reduce costs and streamline the processes.

Methods already exist for conducting business rules validation functions. The Object Management Group (OMG) organization has standardized the use of Semantic Business Vocabulary and Rules (SBVR) for defining a taxonomy and repository of business rules. The International Organization for Standardization (ISO) has defined a standard for asserting validation rules called Schematron for XSLT. OGC has proven through a series of tasks in the OGC interoperability testbed, that service orchestration using the OGC Web Processing Service (WPS) 1.0 standard in SWIM can provide the necessary business processes for validation of SWIM data.

ATM data formats were defined in order to standardize the way in which ATM data producers and consumers exchange information. Adopting standards provided by OMG, ISO, and OGC can reduce the costs of implementation for SWIM technical services such as data validation and improve workflows, ensure data quality, and increase safety.

SEMANTIC VALIDATION CHECKS CONTENT BASED ON BUSINES RULES

Syntactic validation is based on the structural data check as defined by an XML schema or DTD. This check neither includes “business rules” defined by the regional operation of the system using the data, nor it is capable to express some special application dependent constrains. Validation based on business rules means additional rules are specified by domain experts and using well understandable domain language in order to eliminate or reduce the lack of expressiveness in the data structure description taxonomies such as the XML

schema (e.g. cardinalities or data value ranges). Validation is performed by a software component called validation engine.

A full semantics-based approach to data validation is a complex but powerful method in which standard taxonomies of aeronautical data types would be converted into equivalent semantic representation, specified through an equivalent taxonomy/ontology and processed together with accordingly specified business rules. The role of a validation engine would be performed by a computational component called a “reasoner”, responsible for automatic evaluation of data sets, meta-data information, and associated rules in order to identify possible lack of logical consistency in data sets.

BUSINESS RULES ENGINE FOR RULES VALIDATION

In the case of data validation and business rules enforcement, these differ between two distinctive approaches, a procedural and a declarative one. In the procedural approach the encoding of validation rules using an adequate programming language is a step-by-step, iterative validation procedure for each assertion for every validation rule. The declarative approach assumes design of high-level “declarative” statements (i.e. business rules), which represent the data validity requirements or are used to describe the fulfillment of certain business rules.

In the implementation of procedural enforcement, one could program validation classes with a function per assertion, specify rules containing assertions as higher level functions, and finally, compile and pack them into software libraries. As an output, a number of software modules and libraries would be created carrying different sets of rules. These libraries are then deployed into a validation/rule enforcement application server container.

Although there are some benefits in using this method, such as better overall performances and more specificity in cases of very complex rules, the lack of flexibility and low level programmatic approach make this variant rather inappropriate for dynamically changing ***business driven*** data validation and ***business rule*** enforcement. Another major difficulty with this approach is that maintenance complexity requires well-written and well-maintained software documentation.

In the declarative approach, statements are expressed using standardized, high-level description taxonomy (preferably constrained native human language such as SBVR) and stored and managed in a component called a business rule repository. The validation is performed by a dedicated component, responsible for obtaining, parsing and applying these rules. The term, “declarative”, does not mean the opposite of a procedural approach. Instead, it allows domain experts to

describe what the valid aeronautical data are, instead of specifying validation procedures.

Declarative Approach *Describing, what?*	Procedural Approach *Doing, how?*

It is prohibited that **Runway** with assigned **type** equal-to 'RWY' isSituatedAt **AirportHeliport** with assigned **type** equal-to 'HP'	`import org.aixm51.validator.*`
	
	`public static class Validator`
	`{`
	`public Status checkRunway(Runway runway) {`
	`If(runway.type.equal("RWY") &&`
	`(runway.isSituatedAt != null &&`
	`(runway.isSituatedAt.type.equal("HP"))`
	`return new Status(false,`
	`"Runway can't be assigned to heliport.");`
	`else return`
	`new Status(true, "OK");`
	`}`
	`}`

Figure 1 - Declarative versus Procedural Approach (Rules, 2015)

The use of Schematron for XSLT is a concrete example how an XML document might be checked against predefined structural validity rules initially expressed in SBVR. Those rules described conditions beyond those stated in XML Schema, such as value ranges of properties, cardinalities of aeronautical entities, constraints in associations between entities and so on.

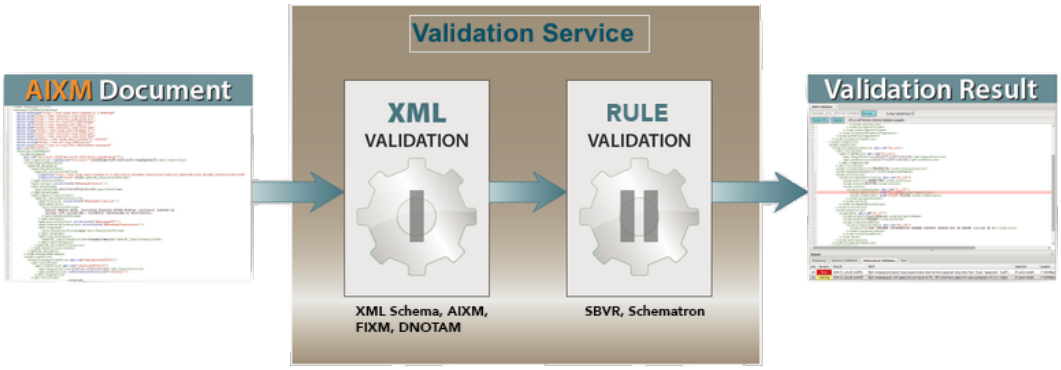

Figure 2 - Validation Workflow

The declarative approach is well suited for business rule enforcement because it provides better support for domain experts to specify high-level rules using a predefined taxonomy using SBVR. The benefit is that all rules are stored in a human readable format for better classification, versioning and monitoring. These rules are then executed by a dedicated software component called a business rules or validator engine using the concepts of Schematron.

OGC WEB SERVICES CAN PROVIDE VALIDATION SERVICES FOR SWIM

In both USA and the Europe SWIM initiatives and corresponding specifications the concept of validation service has partially been specified. In FAA SWIM, there is an infrastructure service (i.e. EMS) classified as one of “common services” containing validation properties, while in the SESAR SWIM similar service has been identified as one of major capabilities of the technical infrastructure. However, it remains unclear, how the validation shall be implemented, deployed and performed considering that both approaches don’t provide exact service specification, capabilities and operational guidelines. It could be assumed that the validation was mentioned just for the sake of completeness and is expected to be performed inside of SWIM compatible services, for example, as part of their service end-points.

During the aviation thread developments of the OGC Interoperability Testbed-11, two tasks were defined to develop a solution for declarative validation based on standardized, high-level validation rules. The first task involved development of an aviation profile for defining business rules using SBVR, and the second task involved translating SBVR to a Schematron for validation purposes. (SBVR, 2014)

The SBVR language was chosen to encode several hundred validation rules beyond those of structural nature already available as part of aviation’s data XML schemas. These rules were then converted into executable Schematron for XSLT format. An XSLT engine was implemented as a validation engine to execute the Schematron XSLT document containing rule assertions encoded in XSLT. This concept has shown good results but needs further improvements in aviation profile for SBVR and conversion to Schematron for XSLT. Current automatic conversion rates are considered to be 60% complete. Further extensions may be needed to support such validation, which would require not only the primary subject of validation (input data) but also complete aeronautical data domain.

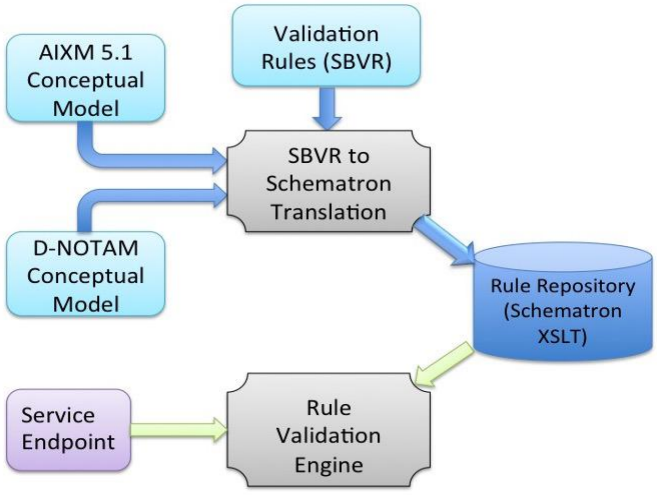

Figure 3 - SBVR-to-Schematron Validation Flow

In the OGC Testbed, the validation service implementation utilizes basic rules sets related to AIXM 5.1, as well as several rules dedicated to Digital NOTAMs. It also performs validation based on AIXM 5.1 XML and several other schema extensions such as airport data, Digital NOTAMS, FAA Digital NOTAM. The service is implemented as a WPS 1.0 compatible service with validation reports encoded using ISO Metadata GMD format. Despite the lack of concrete specifications for SWIM validation, it is proposed that the data validation service for the OGC Testbed be operated as part of SWIM infrastructure deployed and maintained as a cloud service. The service provides a standard access endpoint and supports syntactic schema validation, semantic validation with common purpose business rules, and custom validation with rules per user and use case. Considering the concepts of service orientation (SOA) and cloud computing, such technical service has the following characteristics:

- Centralized deployment in a SWIM cloud environment and available for all authorized aviation stakeholders
- Declarative approach based on validation engine and high level validation/business rules taxonomy
- Predefined common validation and business rules
- Additional validation and business rules per domain or per service user
- The option to invoke the service specifying which validation rule sets shall be considered
- The option to use the service as part of policy enforcement procedure to control service access based on validation results

More information can be found in the publicly released engineering report located at OGC (http://www.opengeospatial.org/projects/initiatives/testbed11) under OGC DOC 15-027 entitled, “OGC Testbed 11 Aviation – Digital NOTAM Validation and Enrichment” (OGC). An example of an OGC standards based validation client can be found at https://swim.m-click.aero/validator/

GLOBAL ATM SYSTEM TRANSFORMATION REQUIRES DATA VALIDATION

The International Civil Aviation Organization (ICAO) has defined needs for validation in certain modules in their Global Harmonization Plan document called the Aviation System Block Upgrades (ASBU) (ICAO ASBU, 2013). In the ASBU document, four Performance Improvement Areas (PIA) are defined, of which PIA2 is the Globally Interoperable Systems and Data. This section defines the needs for SWIM, aeronautical, weather and flight data upgrades as ATM systems continue to evolve. One block upgrade module is Flight and Flow Information for a Collaborative Environment (FF-ICE). FF-ICE Step 1 calls for use of FIXM XML standard formats and lists a series of services foreseen for FF-ICE Step 1 implementation:

a) Initial submission;
b) Validation;
c) GUFI allocation (after the initial flight submission);
d) Nominal trajectory generation (in absence of airspace users defined trajectory);
e) Flight information negotiation (to solve conflict between airspace users' intended flight and existing
f) Flight information update (to change or add to current flight information);
g) Acknowledgement/rejection;
h) Flight information publication;
i) Flight information subscription;
j) Flight information cancellation;
k) Flight suspension; and
l) Flight information. (ICAO ASBU, 2013)

As seen above, the second service listed is validation. This implies beyond simple data validation, as flight plans require a great deal of content validation against the business rules of the FIR for the flight segment being flown. As any pilot or controller can affirm, these rules are complex and change dynamically depending on the region being managed. Within SESAR, "Flight object validation is taking place within the framework of the SESAR projects 10.2.5 and 4.3" and "could be considered as part of SESAR WP/8 and WP/14 in the development of AIRM" (ICAO ASBU, 2013).

VALIDATION IS A CORE TENET OF SWIM

The FAA is conducting phase 2 of the Mini Global SWIM demonstrations in the Florida NextGen Testbed (FTB) with collaboration from several international governments and industry. The Mini-Global Demonstration is a program that tests the exchange of flight, aeronautical and weather information across borders using standard information exchange models" (ICAO MG, 2015). These

demonstrations are seeking to develop a global Enterprise Messaging Service (EMS) network utilizing concepts from industry. "The Mini-Global demonstration focuses on developing seamless information sharing across Flight Information Regions (FIRs), and involves collaboration amongst several ANSPs and operators" (Ngo, 2014).

In order to properly implement data exchange across FIR boundaries and ensure interoperability, the use of data validation is necessary. In fact, core the EMS concept, four tenets must be accomplished by the EMS: Governance, Validation, Security, and Routing (Chen, 2014). While efforts are being conducted within the program to address governance, security, and routing, validation currently only encompasses XML schema validation on XML syntax. Higher-level validation is required.

DATA VALIDATION INITIATIVES IN EUROPE

In Europe, several initiatives regarding the data quality have already been performed providing the documents containing catalogues with guidelines and best practices. Different services deployed in the current aeronautical information landscape in Europe supports certain level of data validation (EAD, NM) but as far as we know, no one reusable, common purpose validation and business rule enforcement service has been deployed and operationally used thus far.

This may change as Eurocontrol begins to implement their Centralized Services initiative, which aims to design, implement and operate a repository of SWIM compatible services. These services provide a centralized solution for the most important AIM operational needs. As part of project SESAR, it has been specified that validation is a SWIM technical infrastructure (TI) service together combined with other essential SWIM based communication enabler service such as messaging, security and monitoring.

In the SWIM-TI specifications there are only high-level formal validation service requirements, without exact endpoint specification and detailed functional description. All three different SWIM-TI Profile specifications (Blue, Purple and Yellow) avoid mandating the explicit use of data validation service. Actually, it is indirectly assumed that any SWIM-TI and ISRM/AIRM compatible service, which fulfills certain requirements and has been deployed in accordance with certain service endpoint specification (AIRM) will automatically fulfill syntactic rules, while deep semantic validation or business rules enforcement remains in the standardization's grey area. However, the lack of formal validation service definition and mandatory use in the SWIM-IT isn't an obstacle to introduce the deep data validation beyond basic syntax checks. Declarative validation service, capable of performing syntactical check, general validation and per user business

rule enforcement, would perfectly fit into the concept of common aviation technical infrastructure (EUROCONTROL, 2014).

BENEFITS TO THE ATM COMMUNITY

The aviation data quality assurance methods described here are based on the declarative data validation and business rules enforcement, tasks which are essential for safety and security. This is a process that should be enforced as a precondition whenever the aeronautical data are created and exchanged or business processes executed. Declarative validation and business rule enforcement has many advantages over ad-hoc programming approaches but it also requires design and implementation of an expert system for access control, rule management and validation execution.

The availability of validation and business rule enforcement services as community-wide available cloud services operated inside or outside of SWIM is the best way to ensure high quality of aeronautical data in technical efficient and costly effective way. Additional benefits are obtained through additional features such as the use case dependent business rule validation using user defined rule sets. Validation results are also applicable for routing and third party service access control - policy enforcement. The service implementation proposed offers full technical infrastructure needed to perform all validation tasks:

- Centralized service in cloud or service oriented architecture
- Standardized service endpoint, user management and access control
- Syntactic data validation
- Semantic data validation and business rules enforcement
- Common purpose and user defined validation and business rules

In the case of many local dedicated implementations per SWIM region, consistency and quality issues combined with high development and maintenance costs are likely to arise. Some of this risk can be avoided through the use of SOA architecture at a global SWIM “cloud” level. It is also crucial that standards such as ISO, OMG, and OGC be adopted for the implementation of data validation services to ensure interoperability and reduce re-work for each SWIM domain user.

CONCLUSION

The aviation data quality assurance methods described here are based on the declarative data validation and business rules enforcement, tasks which are essential for safety and security, whenever the aeronautical data are created and exchanged or business processes executed. This paper provides a recommended approach and proposes technical solutions for data validation and business rule enforcement as part of SWIM technical infrastructure (common service). OGC Testbed 11 implemented a validation service using the approach based on the SBVR encoded validation rules for AIXM 5.1 and D-NOTAMS and Schematron for XSLT as validation engine.

In order to reduce costs and improve reliability and quality of validation it is proposed that a cloud/SOA based, flexible validation platform be implemented as part of SWIM technical infrastructure. The platform should be available to all aviation stakeholders as a service. It should provide basic and extended functionalities for different validation use cases and also supports the use custom validation rules.

Works Cited

Chen, C. (2014, October 01). *World Wide SWIM: Mini Global - A Technical Perspective.* Retrieved September 17, 2015, from SWIM Connect 2014: https://www.faa.gov/nextgen/programs/swim/documentation/media/briefings/SWIM-Connect-2014-World-Wide-Mini-Global-Technical-Perspective.pdf

EUROCONTROL. (2014, June 07). *EUROCONTROL Specifications for Data Quality Requirements.* Retrieved September 18, 2015, from EUROCONTROL: http://www.eurocontrol.int/sites/default/files/publication/files/20140607-dqr-spec-v1.1.pdf

ICAO ASBU. (2013, March 28). *Aviation System Block Upgrades.* Retrieved September 17, 2015, from International Civil Aviation Organization: http://www.icao.int/sustainability/Documents/ASBU.en.Mar.%202013.pdf

ICAO MG. (2015, March 17). *Mini-Global Project Demonstrations.* Retrieved September 17, 2015, from International Civil Aviation Organization: http://www.icao.int/MID/Documents/2015/DGCA%203/IP4_en.pdf

ISO. (2006, June 01). *ISO Publicly Available Standards.* Retrieved September 17, 2015, from International Standards Organization: http://standards.iso.org/ittf/PubliclyAvailableStandards/c040833_ISO_IEC_19757-3_2006(E).zip

Ngo, T. (2014, October 01). *World Wide SWIM: Mini Global Project.* Retrieved September 17, 2015, from SWIM Connect 2014: https://www.faa.gov/nextgen/programs/swim/events/atca_2014/media/ATCA_SWIM_Connect_2014_World_Wide_Mini_Global.pdf

OGC. (n.d.). *OGC Testbed 11.* Retrieved September 17, 2015, from Open Geospatial Consortium: http://www.opengeospatial.org/projects/initiatives/testbed11

OMG. (2015, May 07). *SBVR*. Retrieved September 17, 2015, from Object Management Group: http://www.omg.org/spec/SBVR/

Rules. (2015, September 14). *AIXM 5.1 Business Rules.* Retrieved September 17, 2015, from AIXM Wiki: https://ext.eurocontrol.int/aixmwiki_public/bin/download/Main/AIXM_Business_Rules/AIXM-5.1-BusinessRules-v.0.5.xlsx

SBVR. (2014, June 04). *AIXM 5.1 Business Rules.* Retrieved September 17, 2015, from AIXM Wiki: https://ext.eurocontrol.int/aixmwiki_public/bin/download/Main/AIXM_Business_Rules/AIXM-5.1-BusinessRules-usingSBVRandSchematronV0.3.docx

Aleksandar Balaban

Aleksandar Balaban is an independent ATM consultant located in Berlin, Germany. He received his Graduate Degree in Computer Technology from the Vienna University of Technology. He has been working in the field of aviation for almost 8 years. His experience ranges from software and data engineering to aeronautical information systems definition and implementation utilizing SWIM concepts. In September 2011 he joined the team participating in Europe's most ambitious research and development programme, SESAR. He was working as contributor and task leader in sub projects, which were dealing with system wide information management (SWIM), aeronautical services and Digital NOTAM and Digital Briefing applications. He was also participating in the ongoing work on the OGC PubSub 1.0 project and in the OGC Testbed 11. Aleksandar can be contacted at aleksandar.balaban@m-click.aero.

Charles Chen, CEO – Skymantics, LLC

Charles is a co-founder and CEO of Skymantics, LLC, an aviation consulting firm. He is a subject matter expert in System Wide Information Management (SWIM) research and development with a focus on the Flight Information Exchange Model (FIXM) and Flight Object concepts. He has presented and published in the Air Traffic Control Association (ATCA) Proceedings, Integrated Communications Navigation and Surveillance (I-CNS) Conference Proceedings, and Open Geospatial Consortium (OGC) testbed initiatives. Charles provides SWIM/SOA architecture support for multiple industry testbeds including the OGC Testbed, the FAA Florida NextGen Test Bed (FTB), and the NASA SMART NAS testbed. He is the Aviation thread architect for the OGC Interoperability Program Testbed.