

## *Appendix C: Examples of data standards enhancing “validation” processes and thereby data quality.*

Standardized data enables the use of standardized rules useful in the automation of certain risk assessment, validation, analytical and monitoring processes and procedures across a range of scenarios. In traditional analytical modeling, formulas are typically based upon the physical location of data within a software application’s user interface. For example, a formula that represents a simple concept, such as a ‘current ratio’, would be articulated based upon the physical location of the relevant data (e.g., “B2/F7” in a spreadsheet where B2 is the cell location of “current assets” and F7 is the cell location for “current liabilities”). This physical data orientation hinders the development, sharing, reuse and management of formulas (models) across analytical applications and stakeholder analysts. Machine readable standardized data enables machine readable standardized formulas which can be expressed in a range of machine readable standardized languages including XBRL Formula, RuleML, ISO Schematron, or MathML.

The use of machine readable standardized formulas to test standardized data enables highly reusable and executable analytical formulas or tests of data, based upon the standardized vocabularies. In such a standardized analytical environment, standardized formula or tests are ‘readable’ by both machines as well as business professionals who can share, develop, and collaborate on formulas and tests procedures as they look like a logical sentence (e.g. ‘currentassets/currentliabilities’).

Organizations can use machine readable standardized formulas to express a broad range of validation procedures that test data from reports, ledgers and sub-ledgers. These formulas are very ‘portable’ and sharable, facilitating collaboration and application with appropriate judgment across organizations and disparate ledgers and sub-ledgers within and across organizations. Machine readable standardized formulas or analytical procedures are executable against standardized data from any source systems and thereby enable a more automated and persistent assessment capability. Further, as the machine readable standardized format procedures are reusable across proprietary systems, business professionals can collaborate even more on the governance (e.g. development, maintenance and enhancement) of these standardized format validation and analytical procedures or tests.

Organizations may also use the standardized ledger and related machine readable standardized rules to share and process validation rules by both producing and consuming applications. Organizations use of machine readable standardized validation rules enable execution by both source AND (rather than only) consuming systems; thereby moving validation from consuming to producing systems as a method of enhancing data quality. The FFIEC applied this approach in 2006 and realized a significant enhancement in data quality (e.g. from 66% to 95% compliance)<sup>1</sup>. In the FFIEC case study paper, the collaborative nature of the standardized validation rules enabled application by producing banks as well as the consuming banking regulators. This is clearly visible in **Figure 1**: “FFIEC: Standardized Formula’s Enhances Validation Processes”. The validation rules are presented in both the “Old Process” and “New Process” and in addition to the data quality improvements noted above, the timeline for validation processing improves from 3 days to 10 hours.

---

<sup>1</sup> <http://www.xbrl.org/Business/Regulators/FFIEC-White-Paper-31Jan06.pdf>

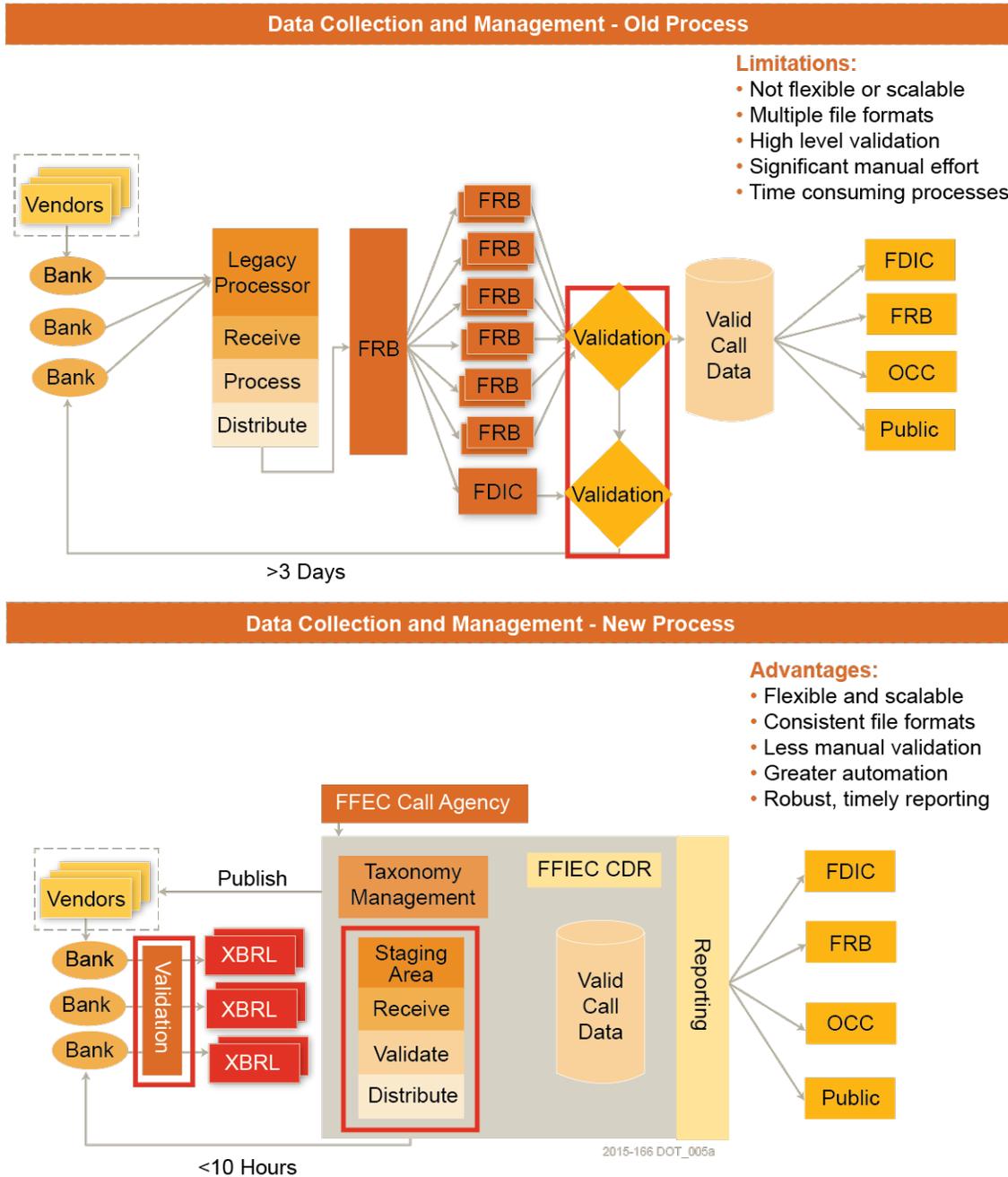


Figure 1. *FFIEC: Standardized Formula's Enhances Validation Processes*

The Security and Exchange Commission's ("SEC") EDGAR Filer Manual<sup>2</sup> provides a range of XML and XBRL related validation rules used by the SEC to validate structured data reports prior to acceptance and can also be used by producers prior to submission. The EDGAR Filer Manual separates validation rules between those for syntax and those for semantics. While syntax can be validated at 100%, semantic validation particularly in subjective areas often requires professional insights and judgment to fully assess data quality and compliance.

<sup>2</sup> <http://www.sec.gov/info/edgar/edmanuals.htm>

**XBRL-CET and BEDES**  
**Buildings University Innovators And Leaders Development (BUILD) - 2015**  
**FOA-0001167**  
Page 3 of 4

The United Kingdom's Her Majesty's Revenue & Customs ("HMRC") Corporation Tax Online Service Validation Rules<sup>3</sup> provides another example of XBRL validation rules that can be applied by both the corporations prior to submission of their data as well as subsequently by the HMRC.

Within the Standard Business Reporting program sponsored by the Dutch Ministry of Finance, standardized validation rules are also available for application by entities prior to submission as well as subsequently by the government. The "Technical Starters Guide XBRL (English)"<sup>4</sup> discusses validation options available when standardized rules are used to enable the application by both data producers and consumers.

The Turkish national tax authority provides standardized rules using Schematron in order to encourage higher quality submissions and move more validation "to the front door". The Schematron rules are an integral part of the submission package<sup>5</sup>, which also includes stylesheets for the necessary human review that will supplement virtually all automated testing.

The [AICPA Audit Data Standards](#)<sup>6</sup> ("ADS") are conceptually very similar to the Intelligent Data standards promulgated by the U.S. Treasury as part of the DATA Act implementation. The ADS is one example of a standardized data model useful to address common challenges that agencies and commercial organizations face is obtaining accurate data in a usable format following a repeatable process. The AICPA has developed voluntary, uniform *Audit Data Standards* that identifies the key information needed for audits and provides a common framework covering: (1) data file definitions and technical specifications, (2) data field definitions and technical specifications, and (3) supplemental questions and data validation routines to help auditors better understand the data and assess its completeness and integrity. The ADS is currently available for General Ledgers and Accounts Receivable Ledgers and other ledgers (e.g. Order to Cash, Procure to Pay, Inventory, Payroll, etc.) are under development and will be freely available upon public release.

Another example of a standardized data model useful to professional managers, accountants, risk managers and auditors is the [OCEG GRC-XML specification](#)<sup>7</sup>. GRC-XML is a family of languages for Governance, Risk, and Compliance information sharing, integration, and communication. It is based on XBRL and XBRL Global Ledger Framework (XBRL GL). GRC-XML has the potential to provide the:

- Basis for an organization to standardize on a common language of Risk and Control;
- Ability to compare the results of risk and control initiatives between organizations; and
- Ability for an organization to integrate information between various GRC systems.

In addition to the specific use of standardized rules, business professionals can also leverage machine readable standardized references or relationships as part of their collaborative knowledge base providing explicit and reusable relationships or links to relationships that may currently be largely implicit. Such relationships can be incorporated into validation processes

---

<sup>3</sup> <http://www.hmrc.gov.uk/softwaredevelopers/ct/ct-onlinevalids.pdf>

<sup>4</sup> [http://www.sbr-nl.nl/fileadmin/SBR/documenten/Technical\\_Starters\\_Guide\\_English\\_march\\_2013.pdf](http://www.sbr-nl.nl/fileadmin/SBR/documenten/Technical_Starters_Guide_English_march_2013.pdf)

<sup>5</sup> <http://www.edeften.gov.tr/dosya/e-DefterPaket.zip>

<sup>6</sup> <http://www.aicpa.org/InterestAreas/FRC/AssuranceAdvisoryServices/Pages/AuditDataStandardWorkingGroup.aspx>

<sup>7</sup> <http://www.oceg.org/resources/grc-xml/#fullcontent>

**XBRL-CET and BEDES**  
**Buildings University Innovators And Leaders Development (BUILD) - 2015**  
**FOA-0001167**  
Page 4 of 4

tying specific data patterns to specific analytical rules as well as explicitly linking unique elements to specific definitions. Standardized references are available within the XBRL Specification as 'Reference Linkbases' and provide a standardized method of logically connecting concepts including:

- USSGL data elements to definitions and/or audit guidance in FAM, FISCAM, Yellow Book and the Green Book, as well as or other relevant tests/procedures, topical subject matter resources, etc.;
- Report level data elements to definitions and/or audit guidance in FAM, FISCAM, the Yellow Book and the Green Book or other relevant tests/procedures, topical subject matter resources, etc.;
- Data scenarios, patterns or risk profiles to specific audit test procedures and assessments;
- Account balances to processes and controls relevant in reconciliation processes and testing assessments;
- Outcomes of validation and analytical test procedures and assessments to additional test procedures and assessments and/or guidance on how business professionals might proceed based upon specific outcomes; and
- Reported data and disclosures to other relevant resources.

As a result of the machine readable standardized format enabling executable validation procedures or analytical tests, which can be applied to different scenarios with appropriate judgment, business professionals can have an enhanced working environment in which validation and analytical procedures are applied in a more automated, persistent and holistic manner while more quickly revealing, sharing, deploying and executing 'best practices' across reports, ledgers and sub-ledgers.

Further, market experience in enabling collaborative processing capabilities indicates that adoption is accelerated among business professionals as they seek to reuse the intellectual property and insights of their peers as a method of improving their specific capabilities and insights. This may be a useful adoption consideration in facilitating viral adoption among supply chain business professionals.

As with other standardized efforts, (e.g. the UPC/barcode), standardized information structures allows for greater levels of automation within business processes thereby further lowering costs, increasing quality and scope of assessments. Some U.S. Federal agencies (e.g. FFIEC, SEC) are already using standardized data and machine readable standardized rules to improve the effectiveness of their analytical processes through automation while others are applying semantically based agents to make subjective assessments on narrative disclosures within standardized data reports.