Implementing XBRL Formulas

Abstract

To ensure the consistency of reported financial data and provide quality assurance of these data to stakeholders, U.S. banking regulators required the standardization of formulas using the eXtensible Business Reporting Language (XBRL). This paper presents the Federal Deposit Insurance Corporation's (FDIC) findings on expressing and implementing XBRL formulas for use with the FFIEC Reports of Condition and Income ("Call Reports").[1] Formulas are applied to data reported by approximately 7,900 FDIC-insured financial institutions to check quality and identify anomalies. This paper[2] discusses (1) the groups of regulatory reporting formulas, as well as their management and exchange between regulatory entities and stakeholders; (2) the need to express formula groups in XBRL syntax; and (3) the implementation of XBRL formulas in advance of an official specification from XBRL International, Inc. (XII).[3]

[1] More information about the Federal Financial Institutions Examination Council is available at http://www.ffiec.gov. A Call Report is used to collect financial information in the form of a balance sheet, an income statement, and supporting schedules. A Call Report is a primary source of financial data used for the supervision and regulation of banks and is relied on as an editing benchmark for many other reports.
[2] This white paper is a draft release and is provided for review and evaluation only.
[3] A formula specification is being developed by XBRL International; see: http://www.xbrl.org.

Table of Contents

Table of Contents 2
1 Introduction 3
1.1 A Standard to Define and Exchange Metadata 4
2 The Need for XBRL-Based Formulas 5
3 Formulas 7
3.1 Consistency Formulas 7
3.2 Characteristic Formulas 7
3.3 Functions 10
3.4 Dependencies 11
3.5 Processing 11
3.5.1 Pre-Submission Processing Model 12
3.5.2 Post-Submission Processing Model 12
3.6 Processing Reduction 13
4 Implementing XBRL Formulas 13
4.1 Vision 14
4.2 Development 14
4.3 Advancement 15
4.4 Software Vendors 15
5 Discussion 15
5.1 Performance 16
5.2 Taxonomy Design 16
5.3 Migrating to a New Formula Specification 17
6 Conclusions 18
6.1 Future Direction 19
7 Acknowledgements 19
8 References 20

1 Introduction

The exchange and validation of financial data are common practices in many industries, and the banking industry is no exception. In fact, FDIC-insured financial institutions (financial institutions or FIs) in the United States and federal banking regulatory agencies (agencies) have been exchanging and validating financial data since the 1800s. Historically, the agencies published data series metadata that included forms and instructions; however, formulas applied by the agencies to check for mathematical and logical relationships were not regularly or systematically provided to the FIs. The agencies experimented with providing formula metadata in text files, PDF files, and spreadsheets. However, the agencies' ability to express precise formulas that described critical relationships within a data series was limited with these formats.

Within the U.S. commercial banking sector, a Call Report is a well known data series that represents a structured and well defined set of quarterly financial data submitted to federal regulators by the majority of institutions. Timely, accurate, and high quality data are critical to regulators and are used in off-site monitoring systems, examination processes, and policymaking. The publication of these data on the Internet promotes financial transparency and helps the FDIC achieve its mission of "preserving and promoting public confidence in the U.S. financial system." The agencies' ability to express formulas that ensure mathematical and logical consistency is critical.

The creation and use of these formulas have evolved over the years. Even though the agencies shared the same common forms and instructions, no common way to communicate and exchange these complex formulas among the agencies, vendors and FIs existed. Eventually each agency developed its own proprietary set of formulas and processing systems, and over time the agencies developed an elaborate process flow that was opaque, inefficient, and labor intensive. The Federal Financial Institutions Examination Council (FFIEC) took the first step toward achieving federal financial regulatory data transparency through the Call Report Modernization project. This effort developed and implemented the Central Data Repository (CDR) which makes use of an international financial reporting standard to define metadata, publish taxonomies, exchange information, and receive and process data reported by federally regulated financial institutions.[4]

The agencies selected XBRL as the non-proprietary, open standard to define and exchange financial reporting data. The use of XBRL would allow FFIEC analysts to define regulatory reporting metadata, such as the concept name (the data element to be collected), data type, instructions, and calculations. However, the functionality to express complex validation formulas was lacking. Alternative solutions to expressing these formulas (such as SQL, Java, and MathML) were considered, but the ultimate solution was to provide an XML layer that would define and publish a formula, provide transparency, encourage re-use, and remain within the bounds of the XBRL specification. The technology needed to include any language that could express a mathematical equation and allow for the "transformation" into another other language, such as SQL, Java, or C#. XPath was selected as a basis for the new XBRL extension because xPath was an XML specification and would provide a way to define "the pointers" to variables that participate in a given formula.

This paper demonstrates the use of XBRL to define and express Call Report formulas with these key findings:

[4] CDR is a new federal financial regulatory reporting collection system developed by the FFIEC. For more information, see: http://www.ffiec.gov/find.

- The XBRL specification can support financial regulatory reporting with the inclusion of a formula specification. XBRL formulas can be successfully exchanged and provide re-use with agencies and other stakeholders.

- Taxonomy design is of high importance and greatly affects the processing and consuming of XBRL formulas.

The contributions of this paper include discussions regarding: (1) groups of regulatory reporting formulas, as well as their management and exchange between regulatory entities and stakeholders; (2) the need to express formula groups in an XBRL syntax; and (3) the implementation of XBRL formulas in advance of an official specification from XBRL International Inc. (XII).[5] The findings provide insight into managing formulas among three regulatory entities, seven software vendors, and 7,900 reporting institutions, and the successful application of XBRL formulas without an official specification from XII. It is important to note that the agencies' implementation of XBRL could not have been successful without the collaborative efforts and business decisions of multiple participants whose primary concern was the management and exchange of financial regulatory data between federal regulatory agencies and report vendors.

1.1 A Standard to Define and Exchange Metadata

Financial regulatory report metadata are classified into four components: financial variables, financial report presentation, financial reporting instructions, and formulas. Prior to Call modernization and the development of the CDR, this information was developed, maintained, and distributed in a variety of formats, including Adobe PDF (for financial report presentation and instructions), Microsoft Word (for financial reporting instructions and formulas), and Microsoft Excel (for financial reporting formulas). XBRL enabled the agencies to create a consistent framework to capture all four components of financial regulatory reporting.

Essential to the FFIEC report framework is a common data dictionary of financial variables, called the Micro Data Reference Manual, or MDRM. Each financial variable defined in the MDRM is defined using a unique number which is used to identify a financial variable on a regulatory report. Multiple reports either make use of or share an MDRM identifier; for this reason, all taxonomy metadata (labels, presentations and formulas) reference the MDRM dictionary. The MDRM dictionary is the foundation to all financial regulatory reporting forms. The MDRM is the 'universe' with instances of the MDRM defined in report taxonomies.

Within the FFIEC report framework, two types of taxonomies exist:

1. **Report taxonomies** are "mapping" schemas that link MDRM numbers with report definitions, such as labels and presentation. In addition, a report taxonomy will contain metadata not found in the MDRM, such as report-specific identifiers, versioning information, and formula message enumerations. Report taxonomies, such as the Call or Y9 series, are "time sensitive" and apply to a specific period of time.

---

[5] A formula specification is being developed by XBRL International; see: http://www.xbrl.org.

2. **Characteristic taxonomies** capture a financial institution's "reportability" which is defined by reporting requirements found explicitly throughout an FFIEC financial report form and implicitly in FFIEC report instructions. These taxonomies do not include presentation or instruction definitions. A characteristic taxonomy is published with each report taxonomy, such as the FFIEC Call or the Federal Reserve Board (FRB) Y9 series. The formulas defined in a characteristic taxonomy are not time sensitive; that is, they are not set to a certain time schedule. A characteristic taxonomy is based on "patterns" found in financial data submissions. These patterns include financial characteristics or events, such as a merger or loss of foreign offices.

The FFIEC report framework was designed with extensibility to other data series. As Figure 1 illustrates, the framework uses a common dictionary which each report and characteristic taxonomy imports. This model provides a modular approach to taxonomy design that can be duplicated and extended to include additional regulatory reports, such as the FRB Y9 series. The FFIEC report framework reflects the CDR data model which uses formulas in both taxonomies to process and validate data received by financial institutions. The same formulas used by the CDR system are used in Call Report vendor software to ensure the transparency of formula results. If a formula processes incorrectly, both the CDR and vendor software should produce an identical result. This approach to pre-validation helps to proactively resolve issues during the report creation and submission process.

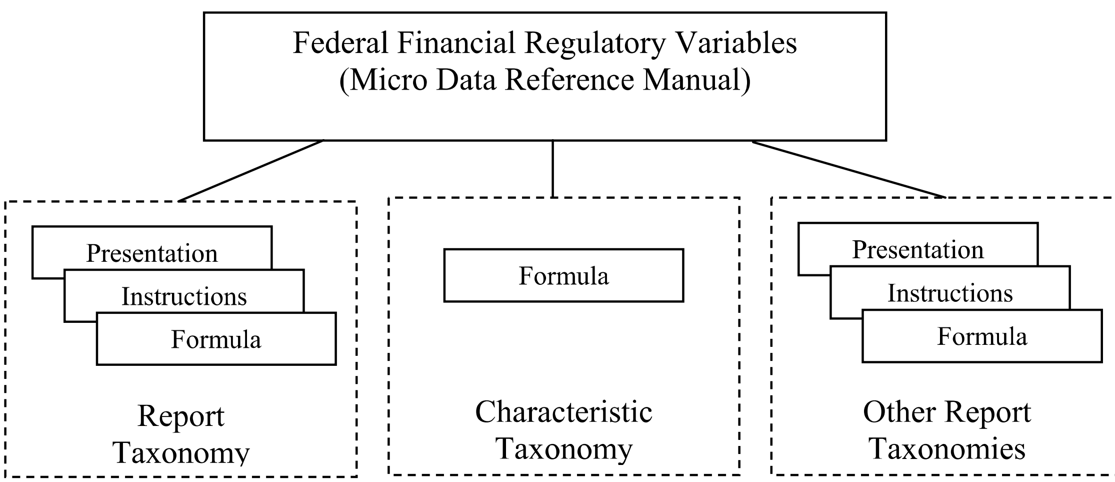

Figure 1 FFIEC Report Framework (conceptual)

2 The Need for XBRL-Based Formulas

The XBRL 2.1 specification provides a data model that successfully captures a large percentage of federal regulatory reporting information, such as terms, data types, presentation information, and report instructions. A majority of the data received by the agencies are validated for consistency and characteristic events, a requirement that the XBRL 2.1 specification only partially fulfilled through its calculation model. XBRL provides a basic validation model that defines calculation relationships among financial variables. Following the XBRL 2.1 specification, a calculation *must* be true or produce an error, but federal financial reporting rules allow financial data to be reported within tolerance levels.

For example, balance sheet items are allowed to grow or contract within defined percentage ranges. If growth rates fall outside of those ranges, the formula(s) will produce an error. The calculation model defined by the 2.1 specification provides a level of simplicity that could not be applied to FFIEC financial data as these data required more complex calculations. The XBRL 2.1 specification does not provide guidance on the use of additional mathematical operands or functions for financial data validation. However, the lack of a formula specification and guidance from XII did not prohibit the agencies from leveraging the current XBRL 2.1 data model.

Early agency pilots produced versions of the FFIEC report framework that included a formula taxonomy. Pilot results successfully demonstrated that financial data could be validated and produce a pass/fail result by processing XBRL-based formulas. The agencies proceeded to develop XBRL formulas for use with the Call Report Modernization project and established these requirements:

1. **Formulas must be defined in an XML-based language for data exchange.**
   The agencies' decision to use XBRL as the data exchange standard for Call Report Modernization satisfied the XML-based language requirement as the XBRL data model was developed from other XML specifications. What XBRL did not explicitly address was a type of XML language to use in defining a formula. Agency pilot projects defined formulas using XPath syntax which proved limited during our implementation-testing phase.[6]

2. **Formulas must include variables and functions.**
   Agency formulas can reference data in the current period as well as multiple prior periods. Agency formulas include attributes such as "-P1Q" or "-P1Y/6-end," which define "prior year" and "prior year and end of June," respectively, which allow references to the same variable across many time periods. Agency formulas include pre-binding functions such as "ExistingNonNil," which returns a result (1) based on the pre-binding value of a reported item or (2) only if the item has a reported value regardless of the XBRL nillable attribute.

3. **Formula process model must execute only with reported data.**
   The agencies discovered that expressing formulas with mathematical attributes provided only a partial solution. A formula process model needed to be developed to process formula attributes and functions with reported data. Prior to CDR, formula processing would operate in "batch mode." This process executed all 2,500 formulas regardless of whether they applied to the reported data. The agencies developed a process model that executes formulas only when a financial variable is present in the reported data. The formula will not execute if a financial variable participating in the formula expression is not present in the reported data.

4. **Formula process model must support cascading data.**
   Agency formulas capture the reportability aspect of a financial institution in a formula type called "characteristic." Characteristic formulas reference multiple data inputs and provide systematic results of the referenced data. That is, data processed with characteristic formulas are "cascaded" to produce a known result. Prior to CDR, FFIEC reporting requirements for financial institutions were implicitly defined in a Call Report instruction booklet. The agencies developed a process model that explicitly defines reportability and ensures that all required financial variables are reported.

[6] xPATH is a language to match and extract parts of a XML document; see: http://www.w3.org/TR/xpath. Also see "Section 4 – Implementing XBRL Formulas."

3 Formulas

Call Report formulas are a shared set of mathematical expressions that check for logical inconsistencies in data received by financial institutions. Formulas can be grouped in two ways -- by consistency or by characteristic. Consistency formulas check for basic accounting errors, such as math, yield, or fluctuations that affect current or multiple periods of reported data. Characteristic formulas, "event driven formula[s] that create a characteristic," include which variables should be reported by a financial institution based on a current event or a prior characteristic, such as "Participated in a Merger" or "$300 Million in Assets."

3.1 Consistency Formulas

The FFIEC Call Report consists of two primary data collections: a Domestic/Foreign collection and a Domestic-Only collection. Each collection has an associated calendar quarter date, such as March 2006. Each collection consists of one or more report contexts, such as "Changes in Equity Capital." Certain report contexts are required during specific quarters, such as "December Only" or "Financial Variables Reported in June Only." Consistency formulas are expressed to validate financial variables in a single context, such as "Changes in Equity Capital," or across multiple contexts, such as "Net Income from Changes in Equity Capital Must Equal Net Income on the Income Statement."

Report Context	Consistency Formula (Relational)
Single	"Charge-offs" on "Loans secured by real estate to non-U.S. addressees" must be less than or equal to the sum of the components of "Charge-offs" on "Loans secured by real estate"
Multiple	For securitized assets that are outstanding as of the report date, the related charge-off items are reported on a calendar year-to-date basis. Therefore, the amount your bank reported this quarter for "Charge-offs" sold and securitized should be greater than or equal to the amount reported in the previous quarter.

Table 1 Consistency Formula (relational)

Table 1 shows an example of two consistency formulas categorized as "Relational." A relational formula compares a financial variable with peer variables within or across different report contexts and must either equal, be greater than, or less than the compared item. Relational formulas will include one or more periods of data. The agencies have categorized Consistency formulas into eight groups: Actual/Null, Relational, Equalities, Itemization, Yield, Ratio, Zero-Value Testing, and Text/Non-Financial. (For more explanation of each of these groups, see "Appendix - Consistency Formula Detail.")

3.2 Characteristic Formulas

Federally regulated financial institutions follow a set of reporting requirements as defined by the FFIEC. FFIEC reporting requirements are summarized into four categories: who must report, what must be reported, where the report must be filed, and when the report must be received. These categories or 'patterns' are captured and expressed in characteristic formulas to ensure that a financial institution is reporting the correct data based on previously captured bank attributes, such as asset size, or based on current or prior period events, such as a merger/acquisition.

A financial institution falls into two primary patterns, either Domestic-Foreign-Base or Foreign-Only-Base. Additional patterns are possible depending on the institution's characteristics for a given reporting cycle (see Table 2 for a sample list of Call Report patterns).

Call Reporting Patterns
Domestic-Foreign-Base
Foreign-Only-Base
Foreign-Only-Reporting-Status-Medium
Foreign-Only-Reporting-Status-Large
Domestic-Foreign-March-Always-Only
Foreign-Only-March-Always-Only
December-Always-Only
June-Always-Only

Table 2 Call Report Patterns

Call Report patterns are applied to financial institution data using a unique process called a *cascading formula pipeline* [7]. In this process, characteristic formulas have two parts: (1) patterns are defined as reportability rules and (2) 'filter' formulas called reportability edits are applied to various stages of submitted financial data. The first step in this process occurs when a financial institution applies reportability rules to a set of data or reportability instance. These data are the result of a set of questions which requests information on prior quarter assets, bank characteristics, and current period events from a financial institution. The data contained in a reportability instance are not an FFIEC reporting requirement and are not submitted by financial institutions in the Call Report. However, in a cascading formula pipeline, reportability instances are *required* and must be available for processing by reportability rules (see Table 3 for a sample set of reportability questions and answers).

Reportability Question	Answers
DURING THE CALENDAR QUARTER, DID THE INSTITUTION ACQUIRE ASSETS OR LIABILITIES THROUGH A BUSINESS COMBINATION OR BRANCH ACQUISITION, OR DID THE INSTITUTION COMMENCE BUSINESS AS A NEW INSTITUTION?	No
AT ANY TIME DURING THE CALENDAR YEAR, DID THE INSTITUTION HAVE AN INTERNATIONAL BANKING FACILITY (IBF) ESTABLISHED IN ACCORDANCE WITH THE TERMS OF FEDERAL REGULATION D?	No
ANY TIME DURING CALENDAR YEAR, DID THE INSTITUTION HAVE AN EDGE OR AGREEMENT CORPORATION ORGANIZED UNDER SECTION 25 OF THE FEDERAL RESERVE ACT AND SUBJECT TO FEDERAL REGULATION K?	No
WHAT IS THE NUMBER OF EDGE OR AGREEMENT CORPORATIONS OWNED BY THE INSTITUTION AS OF THE REPORT DATE?	0

Table 3 Reportability Questions

[7] Cascading formula pipeline is a set of sequential processes a formula processor follows in order to present a desired result. Each subsequent result can be input for the next process, until the desired result is reached. Formulae act as a filter to constrain data through the pipeline.

Characteristic formulas are designed to execute sequentially with instance data. The first step in the sequential processing is to provide an input document. A set of formulas is then applied to the data and a validation process produces an intermediate result. A second set of formulas is then applied to the intermediate result, and a second validation process produces another result. Figure 2 illustrates a financial institution's reportability data (the input document) producing an intermediate result. In this example, the processing of reportability rules has produced an intermediate result that matches a medium-sized financial institution reporting for domestic offices only.

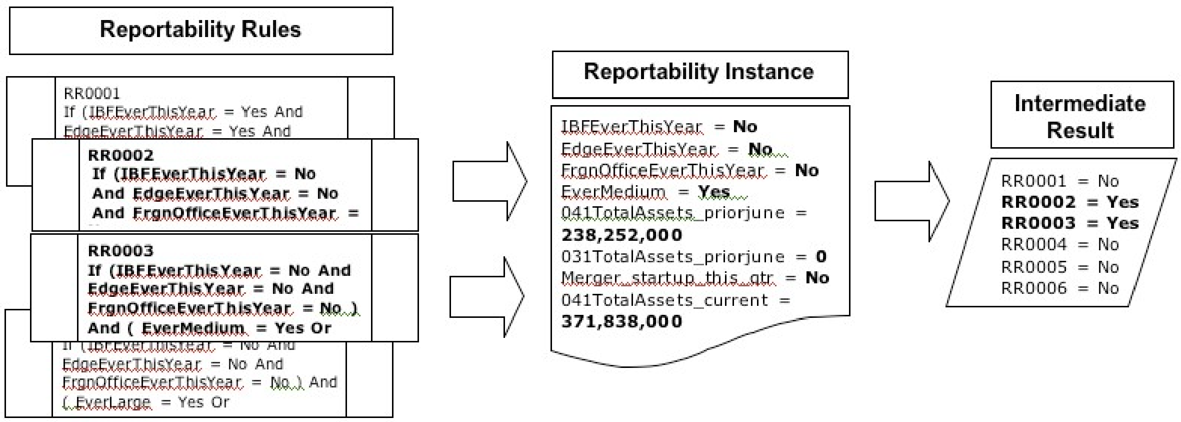

Figure 2 Sequential Processing, Step 1 – Formula to Data and Result

The result of processing reportability rules with data is considered "intermediate" because additional variables must be processed in order to produce the desired result - what financial variables are required to be reported by a financial institution. A second set of characteristic formulas called reportability edits or 'filters' are applied to the intermediate result. Reportability edits contain references to every financial variable in a given financial report, as well as each reportability rule. When processed with the intermediate result, these 'filters' perform a check to determine which financial variable is required to be reported by the financial institution. Figure 3 illustrates processing of reportability edits and the intermediate result to produce a second final result.

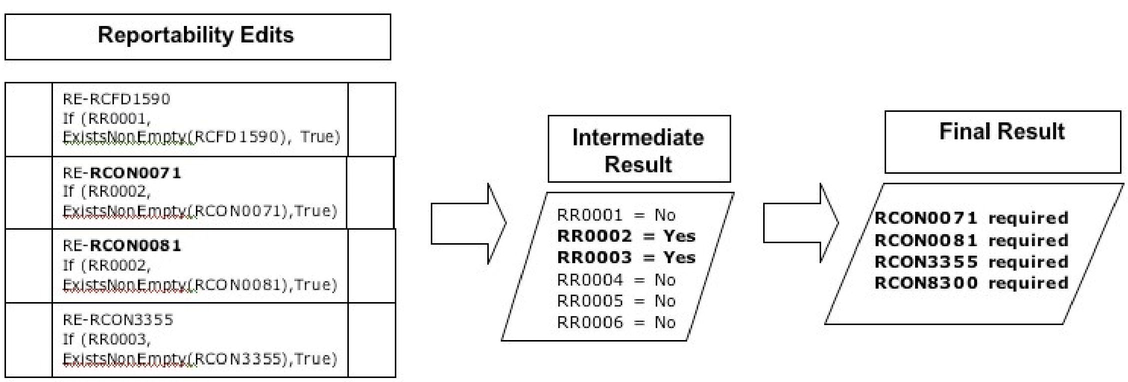

Figure 3 Sequential Processing, Step 2 – Edits to Data and Result

To better understand how a list of required financial variables is provided using characteristic formulas and reportability data, we must understand the meanings used by convention for the value "False" (no) in different kinds of formulas. Figures 2 and 3 not only illustrate the sequential nature of processing characteristic formulas, but also that reportability rules and edits respond differently to data that are "False." When reportability input data contain "False" (no) and rules are applied, the result is captured in the intermediate data output and provided for edit processing. If the intermediate data contain a "True" (yes) value when edits are applied, the result is considered "passed" and the formula processor continues on to the next formula. When the intermediate data contain a "False" and edits are applied, the result returns an error message that states: "Variable xxxx is required." Figure 4 show the decision made by a reportability edit depends on the state of the intermediate data.

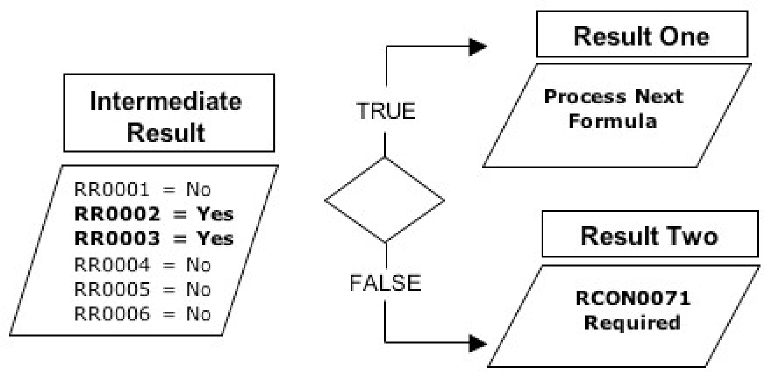

Figure 4 Reportability Edit Decision Logic

3.3 Functions

Explicit validation rules that cannot be defined as an XML data type are captured in special functions referenced in consistency and characteristic expressions.

Table 4 explains two functions used in FFIEC formulas. The *ExistingOf* function in Example A will result to "True" if all variables return a value. If some or no values are returned, the formula processor will not process the formula and will proceed to the next formula. In business terms, Example A states: "Did the bank, at any time during the calendar year, have an international banking facility or an Agreement Corporation or any foreign offices, Yes or No?" The *ExistingNonEmpty* function in example B will return "True" if the financial variable exists in the submitted data for the report and is not nil. In business terms, Example B states: "If bank is reporting with Domestic Offices Only then Current balance from Foreign or Central Banks must be reported." *ExistingOf* and *ExistingNonEmpty* are two functions from a larger set of functions expressed in FFIEC formulas.[8]

[8] See Section 2.10 of the *CDR Interchange Specification 1.3* for a complete list of custom FFIEC functions.

A	ExistingOf	(ExistingOf(cc:RCONC587[P0], False) = True Or (ExistingOf(cc:RCONC588[P0], False) = True) Or (ExistingOf(cc:RCONC590[P0], False) = True))
B	ExistingNonEmpty	If (rr:RR0002.ExistsNonEmpty(cc:RCFD0074[P0]), True)

Table 4 *ExistingOf* and *ExistingNonEmpty* Functions

3.4 Dependencies

Certain formula functions depend on the results of other formulas and are primarily used in processing characteristic data. For example, RR0002 is considered a variable in Example B. If RR0002 is true and RCFD0074 exists, then the formula processor proceeds to the next formula; if RR0002 is false, then the formula processor returns an error message. XBRL does not provide any guidance on processing formulas or data with dependencies from other sources. The agencies developed their own formula process model to accommodate business-specific functions and processing.

3.5 Processing

Characteristic formulas are expressed to handle two processing models, pre and post. Then agencies developed a custom formula processor to handle both pre- and post-processing of XBRL formulas. These processing requirements were implemented using custom functions, such as *ExistingNonNil.*

Characteristic and consistency formulas follow different processing models. Consistency expressions are defined to process data and provide a result. Characteristic expressions are defined to process data, provide a result, process the result, and provide a second and final result. This type of "cascading" data processing is a critical step to understand how financial data are processed in CDR. Validation must follow a fixed order of execution to provide a proper result. Figures 5 and 6 illustrate an overview of the cascading formula pipeline used by CDR.

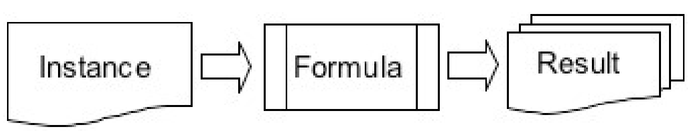

Figure 5 Cascading Formula Pipeline – Consistency

Consistency formulas follow a basic process: formulas applied to input data produce result messages (see Figure 5). Starting with input data, the processor applies a formula and returns a result message *only* if the formula returns the value "false."

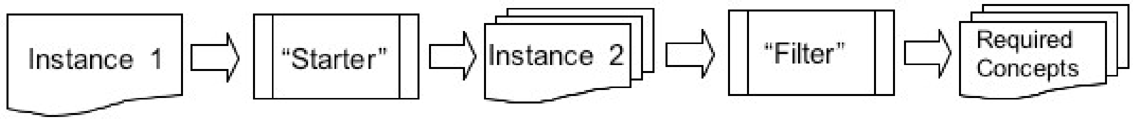

Figure 6 Cascading Formula Pipeline – Characteristic

By contrast, characteristic formulas are designed to execute in a fixed order with a processor executing a "starter" set of formulas called "reportability rules." Reportability rules are applied to financial institution reportability data which return an intermediate result. A second set of "filter" formulas, called "reportability edits," are applied to the intermediate data, which return a result such as "Bank Has Domestic Offices Only with Trading Accounts." The result is a list of variables a financial institution must report for a given Call Report period.

3.5.1 Pre-Submission Processing Model

Figure 7 details the pre-submission model with parentheses delineating a process in vendor software not visible to a financial institution.

1. Set of reportability questions proposed to financial institution
2. Institution answers reportability questions (a process creates data for answers)
3. (Rules applied to reportability instance)
4. (Intermediate result is returned)
5. (Edits are applied to the intermediate result)
6. List of required variables are returned
7. (Software may "gray-out" or not include certain variables in financial report software.)
8. Report data filled out by financial institution
9. (Report instance created and consistency formulas are applied)
10. Validation results are provided to financial institution (via software)

Figure 7 Pre-Submission of Formulas

In pre-submission, the overall outcome is to provide a list of variables that must be reported by a financial institution. Vendor software used by the financial institution then has the option of presenting only the variables that must be reported based on reportability edit results. The financial institution inputs data into the software and then applies consistency formulas which result in failure messages if data validation fails.

3.5.2 Post-Submission Processing Model

Validation performed by the agencies before data are stored in CDR based on the post-submission model. The input document to start validation is reported data, as opposed to a reportability instance. Steps to produce data validation are slightly different, but the processing of formulas is the same and must follow a set order to provide the correct result. Figure 8 describes the post-submission model with parentheses delineating a process in CDR.

1. CDR receives a report instance from financial institution
2. (CDR identifies reporting institution and returns prior period reportability data)
3. (CDR creates composite of a report instance and a reportability instance)
4. (CDR applies characteristic and consistency formulas to composite instance)
5. Validation returns a list of required variables and any possible consistency failures

Figure 8 Post-Submission of Formulas

In post-submission, the result is to provide a notification to the financial institution with possible errors. When the CDR receives a report instance, a process identifies the reporting institution and retrieves prior-period reportability data. CDR creates a composite instance of the reported data and reportability data. Characteristic and consistency formulas are applied sequentially to the composite instance (first characteristic then consistency).

3.6 Processing Reduction

The CDR formula processor applies consistency formulas only to data reported in a report instance. That is, if a financial institution does not report a set of variables, the processor does not execute formulas that contain the missing variables.

Characteristic formulas are designed to ensure that a financial institution is reporting its minimum requirements, but these formulas do not preclude a financial institution from reporting more information than is required. If a reporting institution reports more data than is required, by reportability rule standards, those data are subject to additional consistency formula processing. Figure 9 sets forth the execution process of formulas and data. Formulas 1, 2, and 4 are applied to the report instance, and Formula 3 is ignored.

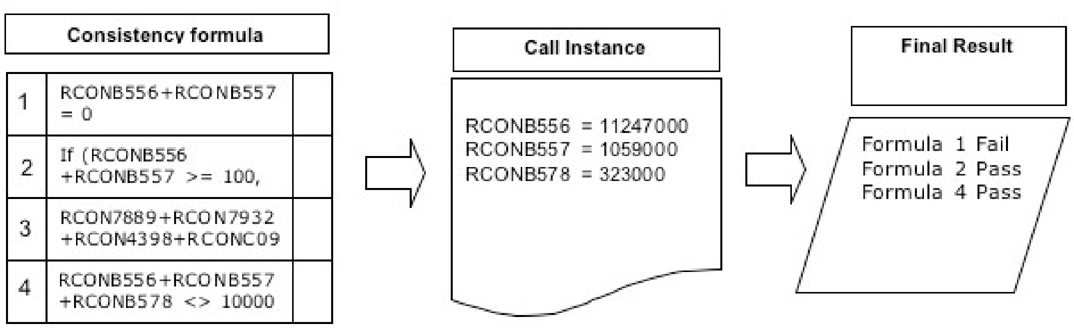

Figure 9 Application of Formula Processor

4 Implementing XBRL Formulas

The FDIC has been involved with the XBRL consortium since 2000 and became an official member in

2001. It was an early adopter of the XBRL specification and started taxonomy development work using XBRL v1. The first taxonomies developed using XBRL v1 were the FRB Micro Data Reference Manual (concepts.xsd) and the FRB National Information Center (NIC), which contains reportability data for all financial institutions. These early taxonomies defined metadata and concept relationships successfully but did not provide a good solution for XBRL formulas. Early drafts of the taxonomy frameworks included a template for XBRL formulas with formula names and definitions. However, formula expressions were not included. XBRL v1 schema and calculations did not provide the functionality needed for formulas.

The release of XBRL v2, which included XML technology for linking schema resources, provided the agencies with the functionality to create a formula taxonomy. The new XBRL v2 specification provided a better taxonomy design model that enabled the agencies to create a modular taxonomy framework with a common taxonomy schema. The Federal Reserve's MDRM, a dictionary of shared data names, has become the foundation for all other FFIEC report taxonomies. XBRL v2 linkbases allowed the agencies to create custom linkbase files for report instructions and formula messages (based on the label linkbase). XBRL v2 enabled the agencies to create a presentation solution for representing a financial reporting form with rows and columns (a combination of presentation and reference linkbases). Most importantly, XBRL v2 provided functionality to design and create XBRL-based formulas.

4.1 Vision

The FDIC conducted a pilot project that provided the foundation for the current CDR system. The pilot project included a proof of concept (POC) that was essentially CDR in a nutshell. The POC included a working demo of a Call Report processing application that resided in an Access database. The demo application included six processes: Legacy Data Import, Taxonomy Creation/Distribution, Software Vendor Simulation, XBRL Instance Submission, XBRL Instance Validation/Store, and Data Distribution. An XBRL v2 Call Report taxonomy, which included a formula taxonomy, was developed for the POC demo. The report and formula taxonomies were fully functional and successfully validated Call Report instances with formula result messages. The POC formula taxonomy consisted of a formula schema which defined certain formula expression attributes such as formulaLink and formulaArc. The formula taxonomy included a formula taxonomy schema which defined FFIEC-specific attributes, such as formula names and formula message enumerations. The formula taxonomy included two linkbase files, one for formula expressions and one for formula messages. The formula expressions linkbase used XPath syntax to express consistency formulas and included global declarations, such as "P0" for the current period and "-Q1" or "-PY1" for prior-period offsets. The formula messages linkbase included text messages for formula failure messages, such as "Assets this quarter must be greater that last quarter. Please review data and resubmit." In addition, XSL style sheets were used to validate Call Report instances with formulas.

4.2 Development

The FDIC developed and enhanced the FFIEC taxonomy framework and initiated development on the XBRL Business Analyst Tool (xBAT). This tool marked the beginning of a full CDR implementation using XBRL as the exchange mechanism and brought the POC demo into reality. The Call Report and formula taxonomy design remained virtually the same but included absolute and relative context references. Call Report taxonomies are published on a calendar quarter. However, formula expressions reference prior period data, and a formula processor will need a point of reference when processing data with formulas. The xBAT formula taxonomy included absolute context definitions, such as “P0” for current period or “P1” for one period prior. The xBAT formula taxonomy also included relative contexts, such as “–P1Y” for the prior year or “–P1Q” for the prior year quarter. The xBAT formula taxonomy was a simple implementation of XBRL formulas and did not contain any special functions or processes, such as reportability. The xBAT formulas followed a simplistic implementation of cascading data validation where formulas process report data and provide a result message. Formula expressions did not require a special or custom processor to process formulas with data, but the final release of xBAT did provide a mechanism to validate report or formula taxonomies with instance data. The formula design was sufficiently simple so that any off-the-shelf XML processor could be used to process the formulas with instance data. Also, xBAT provided Call Report formulas in a separate formula taxonomy. This allowed a software vendor to process formulas with data without having to consider the report taxonomy based on financial reporting forms.

4.3 Advancement

Building on the foundation work produced by the POC demo and xBAT, the Metadata Management Tool (MMT) was developed to maintain and publish Call report taxonomies for CDR. MMT introduced major changes to the existing FFIEC taxonomy framework. The taxonomies were updated from XBRL v2 to v2.1 and introduced taxonomy design changes. The FFIEC taxonomy framework included report taxonomy with formulas and characteristics taxonomy with formulas. The most important changes were the removal of a formula taxonomy and the inclusion of a formula linkbase in the report taxonomy. These changes placed the formulas at the level of report presentation, and formula definitions are now included with financial report presentation variables. MMT also introduced a characteristics taxonomy which allowed the agencies to verify that a financial institution is reporting the correct data.

4.4 Software Vendors

In 2004, the agencies began exchanging XBRL taxonomies with Call Report software vendors. Call Report software vendors became the first non-XBRL software companies to consume and develop software based on XBRL taxonomies. The agencies started weekly communication with software vendors and provided ongoing assistance. As development continued, vendor questions became more focused in the areas of schema validation, reporting null variables, whether multiple instances could be submitted as one, and if numbers should be reported whole or in thousands. The XML and XBRL specifications did not provide guidance in certain situations, and other situations did not match agency business practices. The agencies created the CDR Interchange Specification which provided additional guidance for software vendors and added a business-specific layer to the XBRL 2.1 specification to address agency technical and business requirements.

5 Discussion

XBRL formulas can be used successfully in government and corporate projects to ensure data consistency and improve business processes and transparency through increased pre-validation of reported data. However, certain issues arise when implementing XBRL formulas on a large-scale enterprise project. Although it is not the purpose of this paper to discuss implementation issues, four areas of concern have been identified that must be considered for a system using XBRL formulas. Areas of concern may be classified as project specific, and some can have an adverse effect on a formula implementation. The agencies are more advanced than other U.S. organizations in implementing XBRL, and consequently agencies' experience is strongly influencing the XBRL draft specifications. In both cases, XBRL formulas work and must be considered for any financial regulatory entity or auditor that ensures consistency for financial data. Each issue was either an effect of, or had an effect on, the formula decision and design. Issues encountered include: (1) performance, (2) taxonomy design, and (3) migrating to new formula specifications.

5.1 Performance

The agencies have implemented an Internet-based system for collecting and processing financial regulatory reports submitted by approximately 7,900 financial institutions. Financial data submitted to the system are validated with XBRL formulas immediately upon receipt, and an e-mail notification with possible validation failures is sent to the submitter. The CDR system does not translate the XBRL formula expressions into another format, such as SQL or Java script. Instead, the CDR system uses an XBRL processor (UBMatrix Automator) to process XBRL formula expressions with financial data. The natural state of using XBRL and not SQL code came into question when CDR experienced slow data processing response time.

Formula processing using XBRL is slower than when using mainframe or SQL code. This became evident when the FDIC moved from a Call Report processing mainframe environment to the new Internet-based environment. CDR experienced performance problems in late pre-production tests. Pre-production performance problems were based on taxonomy design, particularly when the formula was included in the report taxonomy as opposed to a separate formula taxonomy. XBRL-aware processors follow Discovery Taxonomy Set (DTS) rules which follow a set path when processing a taxonomy file.[9] That is, the DTS will discover any schema or linkbase references defined in a given instance and proceed to load the information in memory. When an instance is processed in CDR, the processor then loads the report taxonomy which includes not only the formula but also report instructions and the report presentation. The discovery of a complete report taxonomy during formula processing adds time, resulting in slow performance. The issue was compounded if multiple prior-period amendments were resubmitted or a single prior-period validation occurred. The issue was addressed by caching current-period and prior-period taxonomies.

Performance issues have been substantially addressed and improved and should not be a significant issue for similar future implementations. Additionally, the benefits of having sophisticated rules abstracted from specific systems improve the portability and transparency of the metadata as well as the maintainability of the system.

5.2 Taxonomy Design

The FFIEC taxonomy design has been through many development iterations, but only two designs have been available for public use. The first taxonomy framework included a report taxonomy and a formula

---

[9] "Overview of XBRL Instances" http://www.xbrl.org/SpecRecommendations/.

taxonomy. The report taxonomy captured Call Report form attributes such as rows/columns, line-number labels, and reporting instructions. The report taxonomy was designed to preserve the layout of a paper-based financial report. The formula taxonomy captured consistency formulas for a particular data series, such as the Call Report or the Bank Holding Company Report. Formula attributes included formula name identifiers, such as "031/041" or "Y9C/Y9B," and error message enumerations and global constants, such as absolute and relative contexts. The framework design separated report presentation from data validation and allowed a processor to validate data without the overhead contained in a report taxonomy.

The current taxonomy framework design includes a report taxonomy and a characteristics taxonomy. The formula taxonomy has been subsumed within the report taxonomy. With the exception of the formulas, the report taxonomy includes all previous Call Report form attributes. Formula definitions have been moved to the report taxonomy, and formula expressions and error messages are captured in linkbase files. A characteristic taxonomy captures reportability rules and edits for a particular data series. The CDR kept each XBRL instance associated permanently with a single discoverable taxonomy set consisting of schemas and linkbases. XBRL assumes a single validation, processing, and display step; therefore, a multi-step process applied to instances with a single fixed DTS has consequences for performance:

1. Formulas are directly associated with a presentation.
    - When validating Call Report data with formulas, the processor consumes all report presentation metadata along with the formulas.

2. Data validation time is increased.
    - As stated in 1 above, since a processor must first consume the report taxonomy before starting the process of applying a formula with Call report data, processing time is increased.

3. A formula linkbase includes formulas for whole data series (or two related data series).
    - A formula linkbase contains remnants of the formula taxonomy which contains all formulas for a particular data series. Report taxonomies are essentially a filtered context of a data series. That is, the Call Report data series consists of two contexts, a 031-report taxonomy and a 041-report taxonomy. Formulas in a 031-report contain formulas belonging to a 041-report.

4. A special process requires custom functions and a processor.
    - The formula processor must understand business logic and have a cascading formula model which is prevalent in both consistency and characteristic formulas. At this time, additional syntax, custom formulas, and processing are needed in the formula processor to accommodate these.

5.3 Migrating to a New Formula Specification

The agencies have devoted many hours to the design and testing of XBRL formulas. Most of the development time focused on implementing a formula expression language in a new system using an XBRL formula processor. The initial formula design was simplistic and included consistency formulas only. Initial drafts of the agencies' formulas used the xPATH language to express the formulas and did not include special functions. The formula design successfully captured consistency formulas, processed them with Call Report data, and provided a result message. This process was accomplished using off-the-shelf tools such as MSXML and XSLT style sheets.

The initial formula design was developed and reviewed by Call Report software vendors, XII members, and agency staff until the development of CDR. The primary driver behind changing the initial formula design from xPATH to the current design using a proprietary formula language was the need for advanced functions and a processor that could process these functions. Consistency formulas expressed using XPath successfully processed Call Report data, but xPATH could not capture business logic defined in special functions such as *ExistingNonEmpty*. The agencies needed a more robust formula language to express characteristic formulas and a processor that could understand the processing model defined by characteristic formulas. The agencies discussed the possibility of using another formula standard, such as SQL or Java, but decided to use an XBRL processor with a functions library. The second formula design included these special functions and a data model for processing characteristic formulas.

Agency developers spent considerable time developing an XBRL formula language for use with CDR, including the design and testing of new formula logic to ensure the production of correct results. Many hours of training were provided to agency staff and software vendors about how to understand and work with a new formula language. These efforts can be leveraged in the future. However, this assumes the final XBRL formula specification continues (as the public working draft version does) to express the same language, similar functions, and linkbase syntax as found in the Call Report taxonomies.

6 Conclusions

This paper presents the federal banking regulatory agencies' design and implementation of XBRL-based formulas which are likely to strongly influence the final formula recommendation from XBRL International. This paper also illustrates the classifications of formulas and how formulas are managed and exchanged with stakeholders. The need was defined for XBRL-based formulas and the requirements used in developing XBRL formulas; detailed examples of a new data validation model, "cascading formula pipeline," were presented. This paper provides a short CDR "lessons learned" and discusses some areas of consideration when implementing XBRL formulas.

Finally, the federal banking regulatory agencies' experience with implementing an XBRL solution is much broader than the contents of this paper. However, this paper identifies some key observations for implementing XBRL formulas:

- Data must be structured and metadata must be well defined to successfully implement XBRL formulas.
- Differences in processing XBRL formulas as opposed to database code must be considered.
- XBRL formulas must be in an XML language that provides the following requirements:
    - Ability to define functions
    - Ability to define business-layer functions
    - Must provide an XML Info Set or processing model
- XBRL formulas must support transparency for exchange with other government agencies.
- Participate in and influence XBRL groups—voice your requirements and share your experience

for the benefit of all.

6.1 Future Direction

The federal banking regulatory agencies have taken XBRL to the next level with its implementation of XBRL formulas. The agencies have defined a formula process model that uses XBRL formulas to produce different data views. The influence of this model is evident in the consistency and characteristic formulas. Consistency formulas, when applied to data, produce a result message. Characteristic formulas, when applied to data, follow a sequential process of producing “filtered” results. The model followed by characteristic formulas is called “cascading formula pipeline.” The model can be applied to produce unique results in financial reporting and address some common business issues.

The agencies continue to struggle with a shift away from the collection of data based on a financial reporting form. Similar to many government agencies, the federal banking regulatory agencies collect data, and the mechanism of collecting data has traditionally been a paper-based financial form. All data must have a context that assigns meaning, and the financial report form provides this context. The financial report form does not assist in the paper reduction act, nor does it reduce reporting burden. The agencies attempted to address this issue by creating two separate taxonomies, a report taxonomy and a formula taxonomy. The report taxonomy maintains the financial report form and provides a context in which a reporter can understand the data. The formula taxonomy provides expressions used for the validation of financial institution data. The solution was only partial and did not reduce reporting burden. Characteristic formulas were introduced to address the issue of collecting data based on a financial report form.

Characteristic formulas provide a financial institution with a list of required financial variables. The list does not contain all financial variables on the Call Report, but instead includes a filtered set. The processing model followed by characteristic formulas does not have to be limited to required financial variables but can be extended to include presentation variables as well. Characteristic formulas can be expressed to return a report presentation based on the list of required financial variables. Processing XBRL formulas will allow for “dynamic” report presentations based on financial data characteristics. The agencies will pursue dynamic report presentations in future releases of its taxonomy framework.

The use of XBRL formulas to define and capture business rules is a major step toward achieving data transparency across the U.S. government. XBRL formulas can be applied to various aspects of data reporting and collections. In addition to the processes presented herein, XBRL formulas can also be use in:

- Mapping financial variables between two government agencies.
- Extending base or industry taxonomies.
- Filtering a taxonomy to a “light” version.

7 Acknowledgements

The federal banking regulatory agencies’ implementation of XBRL could not have been successful without the collaborative efforts and business decisions of multiple participants who influenced the final XBRL formula solution. The agencies would like to thank XBRL US and International for supporting the Call Modernization project. Special thanks to Charlie Hoffman, David Vun Kannon, Phillip Engle, Yufei Wang, Geoff Shuetrim, Phil Walenga, Steve Hord, and Sunil Matte for all of the late night discussions and sanity checks. We would also like to thank Matasomo Goto, Herman Fischer, Esteban Castorena, Giancarlo Aguilera, and Oxygen for various taxonomy comments and reviews. Finally, we acknowledge and commend the efforts of the seven Call Report software vendors that helped make this project a success: Jack Henry, Fidelity, DBI, IDOM, FRS Global, ITI, and Financial Architects.

8 References

XBRL
Engel, Phillip W., Walter Hamscher, Geoffrey Shuetrim, David vun Kannon and Hugh Wallis. *Extensible Business Reporting Language (XBRL) 2.1.* (2006).
[http://xbrl.org/Specification/XBRL-RECOMMENDATION-2003-12-31+Corrected-Errata-2006-12-18.htm]

XML Schema
World Wide Web Consortium. *XML Schema.* (2008).
[http://www.w3.org/XML/Schema]

XLink
World Wide Web Consortium. *W3C XML Pointer, XML Base and XML Linking.* (2008).
[http://www.w3.org/XML/Linking]

Micro Data Reference Manual
The Federal Reserve Board. *Micro Data Reference Manual.* (2008).
[http://www.federalreserve.gov/reportforms/mdrm/]

CDR Interchange Specification
Montoya, Mark A. and Sunil Matte. *CDR Interchange Specification.* (2007).
[http://www.ffiec.gov/find/taxonomy/index.html]

Improved Business Process Through XBRL
Federal Financial Institutions Examination Council. *Improved Business Process Through XBRL: A Use Case for Business Reporting.* (2006).
[http://www.xbrl.org/us/us/FFIEC%20White%20Paper%2002Feb2006.pdf]