Creating a Standardized Schema for Representing ISO/IEC 17025 Scope of Accreditations in XML Data

Speaker/Author:

David Zajac

Software Engineer, Cal Lab Solutions, Inc.

777 Pettus Rd, Madison, AL 35757

(256) 479-8651

dzajac@callabsolutions.com

Abstract

This paper discusses the progress made on a proof-of-concept effort to develop a standardized way of expressing a calibration laboratory's capabilities and ISO/IEC 17025 Scope of Accreditation (SOA) in a XML format. The goal is to define a standardized schema for representing A2LA, NVLAP and other SOAs in a digital format. Once established, the metrology community can move from manually verifying uncertainties to a more automatable approach. This Metrology.NET® formated XML database can then be used to generate a traditional SOA document. It can also be used to safeguard against laboratories reporting uncertainties that don't comply with their SOA.

Introduction

When a calibration laboratory undergoes an audit for ISO/IEC 17025 accreditation, its capabilities are presented in a document known as the Scope Of Accreditation. The laboratory's capabilities are comprised in an official SOA document carrying a legal signature from the certifying body as a external verification that the laboratory does indeed have these capabilities. The original and existential purpose for an ISO/IEC 17025 Scope of Accreditation is to serve as the deliverable end product of a calibration laboratory's audit. To serve this purpose, a free-text formatted document is perfectly sufficient. However, the information contained in a free formatted document is not machine readable.

The primary use of the information contained in an ISO/IEC 17025 Scope of Accreditation is the legal responsibility that the certified calibration laboratory has, on an on-going basis, to ensure that it does not misrepresent the auditing body's certification in any calibration report that bears the certifying body's logo. This is a requirement that the auditing body places upon the certified calibration laboratory. To meet this requirement the laboratory should have processes in place to check each and every stated uncertainty on every calibration report carrying a certification logo. The laboratory must ensure that all stated uncertainties correspond to the capabilities

section of its current ISO/IEC 17025 Scope of Accreditation. This is a tedious, complex and time consuming process that lends itself to human errors. Numerical calculations, comparisons, and document generation are tasks better suited for a computer and automated data processing.

However, a major hindrance prevents the development of software capable of auto generating fully validated ISO/IEC 17025 calibration reports - the lack of a tool that can intelligently parse a lab's Scope of Accreditation document for the required Calibration Measurement Capability (CMC) data. The lack of an industry standard format for the content and layout of a lab's ISO/IEC 17025 Scope of Accreditation, prevents the development of a standard tool to parse the documents. At present, it is impossible to effectively tackle the all the variations in SOA formats found across a collection of calibration laboratories.

In this paper, we present the progress made on a practical alternative - the generation of Scope of Accreditation documents from data that begins its life in a standards compliant machine readable format. Using this approach, we believe **the insurmountable task of parsing a Scope of Accreditation document is eliminated by the *a priori* existence of its standards compliant source data**. This well formatted source data can then be used to generate the traditional / legacy ISO/IEC 17025 Scope of Accreditation document via templates and software. We believe that first creating a machine readable data source then using that data to generate a human readable document, will prove to be an effective solution to provide access to the required data contained in an SOA document. We will also broach the fact that in spite of the merits of the proposed technical solution, this approach will have to overcome the challenge of being a significant change to established business practices - in that it shifts the burden of the task to a different point in time and potentially to a different set of participants.

Technical Requirements

The Data

The critical data from a Scope of Accreditation that is required to support test report generation is a significant subset of the data required to generate a full Scope of Accreditation document. Complicated provisions for handling a large variety of units of measurement can be largely avoided and none of the formatting data required to layout the document's printed pages is required. The data from a Scope of Accreditation that is needed for test report generation is contained in the Calibration Measurement Capabilities (CMC) section of the document. This data is broadly split into two categories: one being a set of formulas needed to calculate the certified uncertainty limits, the other being a set of metadata (data about data) needed to determine which of the uncertainty calculation formulas is applicable to a given test report value. Both of these broad categories of data can be further broken down.

Each uncertainty formula (which can be as simple as a constant) is expressible as a symbolic algebraic expression, having: a result, a set of coefficients, and a set of variables.   The

symbolic algebraic expression is the foundational required data element. To be usable, each of the algebraic expression's constituent elements must be fully defined. These elements include the result, constants, and input variables. The data required for the result is its quantity type (i.e., length, mass, torque, etc.). The data required for each constant is its quantity type and numerical value. Finally, the data for each input variable is its quantity type and value source.

Note: Only quantity types for the formula's constituent elements are required. Units of measurement are not required because they can be inferred from a quantity type and a mechanism for doing so has been developed that utilizes a supporting database which specifies a single standardized unit of measurement for any given quantity type.

Since the appropriate formula for uncertainty calculation can vary depending upon the answers to the questions: what is being calibrated?, how is it being calibrated?, and what is the nominal value for the test point?, metadata for formula selection that can be compared against the answers to these questions is required. This metadata typically contains criteria for the Device Under Test (DUT) type, the test technique, the equipment being used as standards and/or accessories, and ranges for stimulus values. Since these criteria are typical, special provisions are made to hold their values. In addition, the list of conditions can be open ended, therefore, provisions are also required for holding a boundless list of additional application specific selection criteria.

In contrast to the relatively simple structure of the data required to support test report generation, the data required for SOA document generation contains elements that are highly unstructured. It includes elements needed to specify page margin dimensions, header and footer layouts, fonts, font styles, logo graphics, nested tables, a broad gamut of unit of measurements, numeric formatting specifications, and a host of other page layout and presentation specific control information. Each of the major accreditation entities has their own unique document template and style guides. If Scope of Accreditation documents are to be generated from data rather than keyed in as free form text using a word processing program, all of this must be taken into consideration, and the software that generates the documents must have access to all the data required to layout theses high complexity documents. Fortunately, there is an established and well worn technical base in place for dealing with just this type of problem set, namely, Extensible Markup Language (XML) and its related technologies.

This effort has progressed to the point that a practical solution exists for holding all of the required data for the CMC section of a Scope of Accreditation. Additionally we have developed a practical solution to handle the task of print ready layout and documentation generation for the non-CMC content of any Scope of Accreditation document. We have demonstrated the capability of duplicating several existing SOA document templates. What remains undone is the automated print layout of the CMC section, and the development of user friendly data editors. For now, editing of most of the data required for the project is being done with a commercially available XML focused software product. In addition, a proof-of-principle browser-based editor for the Unit of Measurement database has been created.

XML Databases and Related Technologies

A database can be created using many technologies. Merriam-webster defines a database simply as: "a collection of pieces of information that is organized and used on a computer." So, it would be wrong to assume that a database must have tables with columns and rows. The data we deal with in this effort does not fit well to tabular organization. In addition, as stated earlier, its existential purpose is to contain the data needed to generate a document. Given this, Extensible Markup Language (XML) and related technologies are a perfect solution to how this data can be organized and used on a computer. By choosing XML technologies to hold and process this data, every task required for document generation can follow straightforward and well supported industry standards.

The base of the XML technologies is the XML document; this is where data is held. In addition to the XML documents are standards compliant software products and libraries that can operate on special types of XML documents that have been designed for specific purposes. If an XML document is not one of these special types, it is customary to affix a ".xml" file extension to the document's file name. Each of the special types of XML documents have customary file extensions as well. In the SOA effort, the following XML document types and standard software technologies have been employed:

1) Schemas (*.xsl and *.sch files) - XML has more than one way to specify schemas for data documents. In the SOA effort both rules defining valid data structure (grammar) and rules defining valid data content (business rules) are required and different schema solutions were selected for these two tasks. XML Schema Definition Language (XSD) is used for the task of defining data structure, and the Schematron Language is used for data content. XML Schema Definition Language is a World Wide Web Consortium (W3C) recommendation and the Schematron Language is defined in ISO/IEC 19757-3:2006.

2) Data Transformation Documents (*.xsl files) - Extensible Stylesheet Language Transformations (XSLT) is a common language used for transforming XML source documents into other text file formats. XSLT documents (*.xsl files) are simply special purpose XML files. In the SOA effort, *.xsl files are used as document templates that are merged with XML data from other sources to produce new XML documents. As such, these documents are both data containers (template content) and processing instruction containers (logic for merging in externally contained data). XSLT documents are processed by specialized software products known as XSLT processors.

3) Processing Scripts (*.ant files) - There are a variety of ways of processing XML based files. Another Neat Tool (Ant) is a batch-type script processing language that is implemented within an XML syntax. The Ant interpreter is written in Java, is developed and maintained by the Apache™ Software Foundation, and carries the Apache™ Software Foundation's 2.0 license. Among its many capabilities, Ant has the native ability to move and rename files, to launch XSLT transformations and to launch custom Java code contained in Jar files.

4) Print formatting and Document Generation (*.fo files) - The Apache™ Formatting Objects Processor (FOP) is another Software product written in Java, developed and maintained by the Apache™ Software Foundation, and carries the Apache™ Software Foundation's 2.0 license. The FOP software has the mission of the production of documents formatted for printed form. It uses as source material XML documents that follow the XSL-FO W3C Recommendation. This software tool has the capability to directly generate Portable Document Format (PDF) files that closely match existing SOA documents in layout.

5) Data documents (*.xml files) - Several pure data documents are required for the SOA effort. Several of these are standard external reference data sources whose content is can be cited in the laboratory specific SOA source data file.

Retrieving Data

There are a variety a approaches to retrieving data from XML data sources.   Amongst the simplest methods are XPath and XQuery expressions. Both XSLT and Schematron processes natively utilize XPath and XQuery expressions, and in this effort, all data querying operations require nothing more.

The SOA Document Generation Process

The process for authoring a SOA starts as it does now, with a template provided to the laboratory by the accreditation body. The difference is that instead of using word processing technology, the template for this effort is an XSLT document and the task presented to the laboratory is to author the contents of XML data documents using a special purpose editor rather than filling out a word processing document using a word processing program.

Once the XML files have been authored, the documentation layout and generation processes is fully automated: The contents of the XML documents are merged with the template in a multistep process using XSLT technology until an XML file that conforms with the Formatting Objects schema is produced, and than this file is converted to PDF by the Formatting Object Processor which directly produces the final PDF output file.

The CMC Data Section

The purpose of the data in the CMC section is to hold the metadata required to select CMC uncertainty formulae as well as the full and unambiguous definition of any CMC uncertainty formula.

The CMC data section is contained in a separate file. The name of all elements in the CMC file are prefixed with the characters "unc:" XML uses a element name prefix to define a namespace. In this case the prefix is the first three letters of the word "uncertainty". The purpose of namespaces in XML is to distinguish element names that might otherwise inadvertently be named the same but have more than one schematic definition. This situation can arise when data sources are merged together from separate source documents. Since,in both this effort and follow-on projects, merging data from separate data sources will be common, all XML files used in the effort will use namespace prefixes. For instance there are values cited in the CMC file that are or will be defined in a separate files that get included into the CMC file by reference. These values currently refer to data defined in a separate Unit Of Measurement data file. In the future there will also be citations to terms that are defined in external dictionaries. These future dictionaries will include definitions for the following: Equipment/Accessory types, Test Processes (Techniques), Measurement Categories, and Test Equipment Setup Models. Once in place, restrictions will be enforced that only terms defined in these external dictionaries may be used.

NOTE: The rationale for limiting terms to only those appearing in these dictionaries is the same as the motivation behind the Vocabulary in Metrology (VIM) that is maintained by the Bureau International des Poids et Mesures (BIPM). Many very niche and technology specific terms are used in metrology and it is problematic to assume that the reader of a SOA should be cognizant of them all. By limiting the use of terms to only those found in external openly accessible dictionaries, the SOA reader will be enabled to delve into the definition of any term as needed.

All XML files (just like HTML files) must have a single root element. For the CMC files, the root element is the "CMCs" element The "CMCs" element is merely a container for a collection of "CMC" elements. Each "CMC element holds all the metadata that is selection criteria for an element known as a CMC "Template", as well as all of the data for the template. A unique CMC template is required whenever the CMC formula in combination with the formula's selection criteria change.
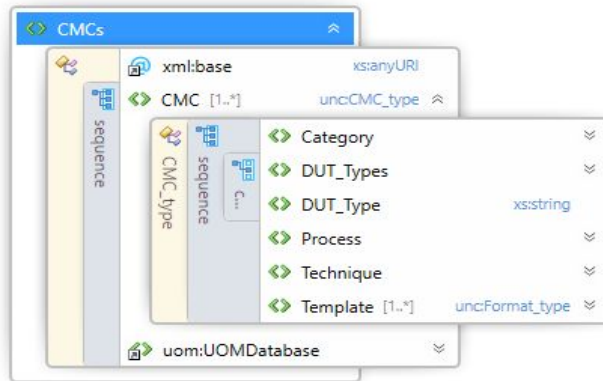
The metadata that is used for selecting a CMC formula tends to exist in a hierarchical structure. The CMC schema recognizes this natural hierarchy and provides for it. A CMC template is a standard first layer of this hierarchy. There are four metadata elements at the CMC template level that serve as the first layer of selection criteria. If any of these four metadata elements fails to describe the subject test, then the CMC template element is not applicable. Each of these metadata elements should in the future refer to terms defined in standard, freely and openly accessible, external XML based dictionaries. These terms are:

1) a broad technical category to which the measurement applies - The category itself can be hierarchical. As an example of a broad measurement technology category, the measurement of a resistor value is likely to occur in a subsection of a metology laboratory devoted to making non-dynamic electrical measurements. A typical category for this type of measurement might therefore be Electrical with a subcategory of Direct Current. In the XML file this would appear as:

```
<unc:Category name="Electrical">
        <unc:Category name="Direct Current"/>
</unc:Category>
```

2) Device(s) Under Test (DUT) type(s) - The Device or Devices Under Test define in coarse terms what is being tested or measured. The DUT type can be a generic description or explicit model/option numbers. If more than one entry is required a "DUT_types" element is required. If however there is only one DUT_type required, the use of a DUT_types element is optional. Therefore in the XML file the DUT type(s) can appear either as a list contained in a DUT_Types element :

```
<unc:DUT_Types>
        <unc:DUT_Type>Fluke 5700</uncDUT_Type>
        <unc:DUT_Type>Fluke 5720</uncDUT_Type>
        <unc:DUT_Type>Fluke 5730</uncDUT_Type>
<unc:DUT_types/>
```

Or as a single entry without the DUT_Types element being required:

<unc:DUT_Type>Fluke 57XX Meter Calibrators</unc:DUT_Type>

3) Test Process Identifier - The Test Process defines what in fine detail on the DUT is being tested or measured. As example, if a plug gage is being tested, possible Process names might be "measure diameter external" or "measure surface roughness". This would appear in the XML file as:

<unc:Process>measure surface roughness</unc:Process>

4) Test Technique Identifier - The Test Technique defines how the test or measurement is performed. As an example, if a gage block is being tested, a possible Technique names might be "non-contacting using laser interferometer" or "contacting substitution comparison". This would appear in the XML file as:

<unc:Technique>contacting substitution comparison</unc:Technique>

NOTE: These four standard metadata criteria (Category, DUT Type, Test Process, and Test Technique) should greatly diminish the need for the "Comments" column found in typical SOA documents. Because the contents of the typical SOA Comments column are nearly unrestricted, they are practically non-usable from a library science point of view (that is, since by definition, the comments columns can contain almost anything, its contents can not not be handled by any practically definable algorithm). A major goal of this effort is to completely eliminate the Comments column from SOA documents.

The CMC "Template" Element

The CMC template contains the definition of the CMC formula or formulae as appropriate, the definitions of all symbols in the formulae, additional selection criteria, and the possible values for all constants defined in the formulae, and selection criteria for the possible values for constants.
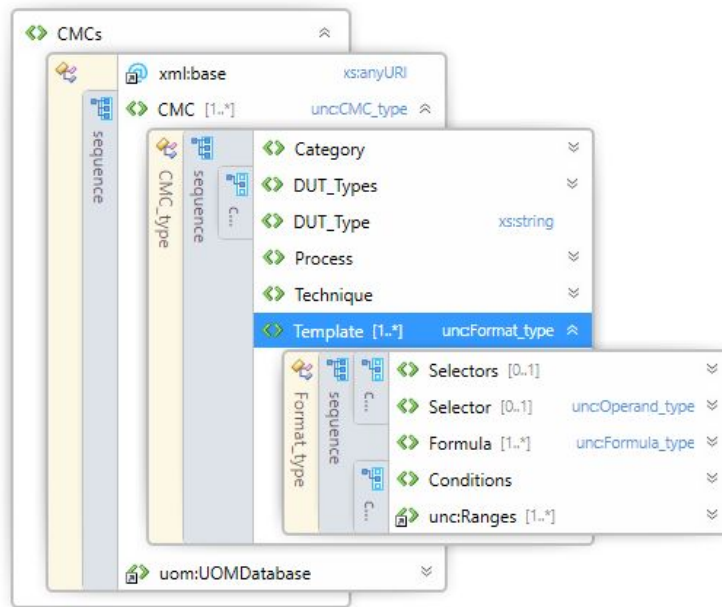
The selection of a CMC formula or formulae set, will often depend of the values of variables. The variables can be numeric or boolean (true/false) conditions. The numeric variables may be explicit in the formula or external to the formula.

There are often symbols that come into play that are not explicit in the formula. These are symbols for variables that act as selection criteria for the template. These symbols are defined in a "Selector" element. The symbol definitions apply to the template level, however, the values that these variables will contain are used by range selection criteria at the formula level. As an example, the CMC uncertainty of a microphone calibration may depend on the frequency of the test point. In this case, a symbol for a variable that will hold the frequency value for different frequency ranges is required and its definition is held in a "Selector" element as follows:

```
<unc:Selector>
    <unc:Symbol>f</unc:Symbol>
    <unc:Quantity>frequency</unc:Quantity>
    <unc:Description>Audio Frequency</unc:Description>
</unc:Selector>
```

In addition to numeric ranges, formula selection can also depend on non-numeric criteria. These criteria can be held in a collection of Condition elements. Condition elements hold boolean name/value pair assertions. This provides an opened-ended method of extending beyond the four standard selection criteria that apply to the template level. If more than one condition exists, they can be organized in either lists or hierarchies. If numeric ranges also exist as selection criteria, they must be contained within the definition of condition criteria. If no condition criteria exist then numeric ranges can exist on their own. This is quite complicated

and a bit more ground work must be presented before an example can be provided. (See Appendix A, CMCs Example for Microwave S_{21} and S_{12} Magnitude and Phase)
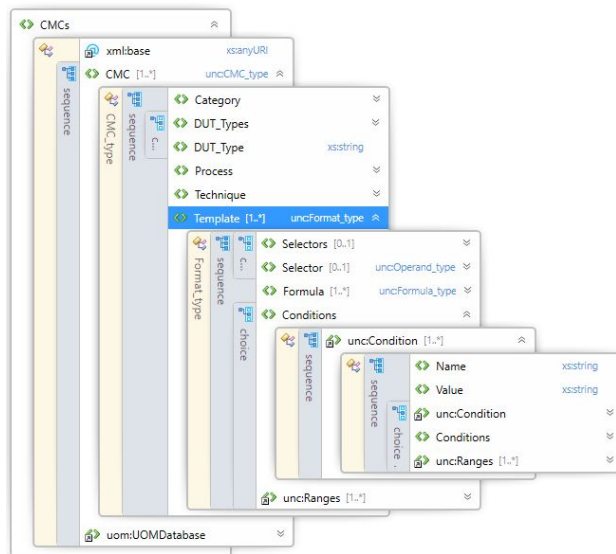

Most of the time, a template will only contain a single formula. There are circumstances however where multiple CMCs apply to a test. A typical case is in the measurement of microwave parameters were one CMC specification exists for the magnitude of the measurement result and a second CMC specification exists for its phase. In this type of case, the default heading name for the CMC column does not suffice because there are two values that must be distinguishable from each other. To handle this situation, the formula element has an optional heading attribute that should be used to override the default column name which is simply "CMC". (See Appendix A, CMCs Example for Microwave S_{21} and S_{12} Magnitude and Phase)
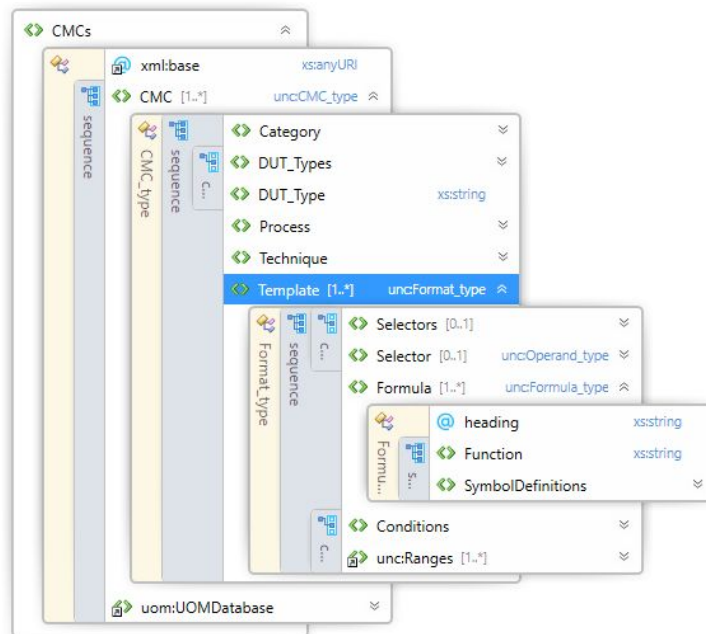
The CMC formula will consist of definable symbols representing inputs and constants, as well as mathematical operators. The formula's expression is contained in a "Function" element. As an example, the uncertainty of a pressure gage often depends on four values: the reading, a percent of reading specification, and a percent of full scale value, and finally the full scale value. In this case the expanded formula's expression (using square root of the sum of squares) might appear in the XML files as:

```
<unc:Function>sqrt((m1 * x + b1 * fs1)^2 +  (m2 * x + b2 * fs2)^2)</unc:Function>
```

This expression is not useful by itself because none of the symbols nor the result has yet to be defined. The schema has provisions to fully define all of the symbols and as well as the result. These definitions are held in a "SymbolDefintions" element that immediately follows the Function element.
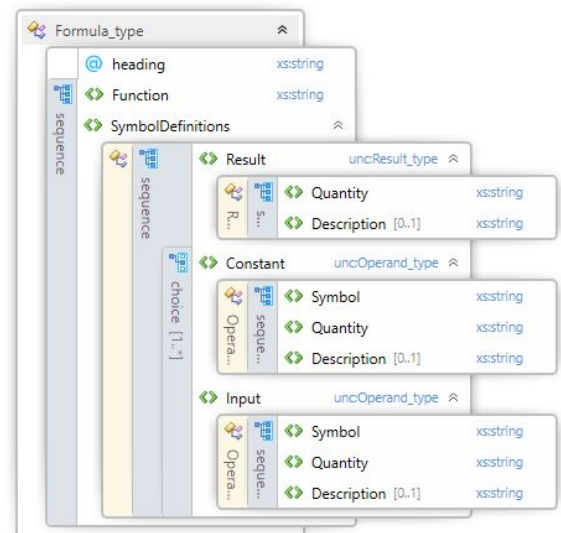
For the above Function, the SymbolDefintions element and sub elements might look like this:

```
<unc:SymbolDefinitions>
    <Result>
        <unc:Quantity>pressure:absolute</unc:Quantity>
    </Result>
```

```xml
<Constant>
        <unc:Symbol>m1</unc:Symbol>
        <unc:Quantity>ratio</unc:Quantity>
        <unc:Description>
                Uncertainty Coefficient of Standard based on Nominal Value
        </unc:Description>
</Constant>
<Constant>
        <unc:Symbol>b1</unc:Symbol>
        <unc:Quantity>ratio</unc:Quantity>
        <unc:Description>
                Uncertainty Coefficient of Standard based on Full Scale
        </unc:Description>
</Constant>
<Constant>
        <unc:Symbol>m2</unc:Symbol>
        <unc:Quantity>ratio</unc:Quantity>
        <unc:Description>
                Uncertainty Coefficient of DUT based on Nominal Value
        </unc:Description>
</Constant>
<Constant>
        <unc:Symbol>b2</unc:Symbol>
        <unc:Quantity>ratio</unc:Quantity>
        <unc:Description>
                Uncertainty Coefficient of DUT based on Full Scale
        </unc:Description>
</Constant>
<Input>
        <unc:Symbol>x</unc:Symbol>
        <unc:Quantity>pressure:absolute</unc:Quantity>
        <unc:Description>Nominal Value</unc:Description>
</Input>
<Input>
        <unc:Symbol>fs1</unc:Symbol>
        <unc:Quantity>pressure:absolute</unc:Quantity>
        <unc:Description>Full Scale of Standard</unc:Description>

</Input>
<Input>
        <unc:Symbol>fs2</unc:Symbol>
        <unc:Quantity>pressure:absolute</unc:Quantity>
        <unc:Description>Full Scale of DUT</unc:Description>
```

```
            </Input>
        </unc:SymbolDefinitions>
```

Non-Technical Hurdles

As a reminder, the overall goal of this effort is to have available, a standards-based database of a laboratory's Certified Measurement Capabilities (CMCs) that can be used in automated processes that generate calibration test reports to assure that said test reports are in compliance with the laboratory's Scope Of Accreditation (SOA). The means of getting to this goal, is to create a means of automatically generating the Scope Of Accreditation document from the desired database, so that the database will exist and be guaranteed, to match the Scope Of Accreditation, because it served as the source for the document.

This effort is at the "Proof of Principle" stage. Enough development has been done to prove that the approach is technically feasible. However, since the accreditation body fully controls the SOA document, the accreditation body must concur with the processes outlined herein. At this point several accreditation bodies have been made aware of this effort and none have chosen to participate in or provide ongoing feedback on its development. As a result, despite the potential benefit to their customer, there is an outstanding concern that no accreditation body will cooperate.

Accreditation bodies do make a best effort to standardize on the format of SOA documents. The International Laboratory Accreditation Cooperation(ILAC) specifies the data that an SOA must contain, and implies a fixed column tabular presentation for the CMC data. In addition, compared to hierarchical tables, tables with fixed columns are relatively easy to create in programs like Microsoft Word™. However, any cursory review of a collection of SOA documents reveals that the data contained in the CMC sections poorly fit the the fixed column tabular format to which the accreditation bodies typically attempt to adhere. The ranges and test conditions that serve as selectors for CMC uncertainties are often more logically represented as hierarchies and are more natural to present in that fashion. In an automated layout process, fixed algorithms must be employed to control layout. Attempts to create layout algorithms that mimic the wide variety of poorly implemented layout approaches currently employed by SOA authors in their attempts to force fit what is naturally hierarchical data into a fixed number of columns would be a poor path to take. Instead, abandoning the the fixed column presentation in favor of a hierarchical presentation would greatly ease the layout problem for a machine algorithm, it would also improve the clarity and standardization of the SOA documents. Again, the accreditation bodies would have to concur with this change.

The ILAC also states that SOAs should be unambiguous. Again, any cursory review of a collection of SOA documents reveals that this requirement is often not met. There are three

main sources of ambiguity in current SOA documents: 1) Terminology, 2) Poorly Defined Formula Symbols and 3) Ambiguous Units of Measurement.

Many terms are used in SOA documents that specific to the nomenclature used within a particular laboratory and there is no requirement that the SOA include definitions for these terms. These terms include items such as DUT type identifiers, measurement technique names, names for standards or accessories, etc. Without including definitions in the SOA, the authors clearly must assume that all potential readers of the SOA document are as familiar with these terms as are the authors. An overlooked and important potential SOA document reader is a customer or potential customer of the laboratory attempting to perform due diligence.

Appendix A - CMCs Example for Microwave S_{21} and S_{12} Magnitude and Phase

In this appendix, an example of a high complexity section from an actual SOA is used to demonstrate several features of both of the word processing created version of the SOA and how its data can be represented following the schema created in this effort.

Sample Extracted from (A2LA Cert. No. 11256.0) 06/09/2015 page 23 of 25

Parameter/Range	CMC[2] (±) Magnitude	CMC[2] (±) Phase	Comments
Attenuation – Measure, S_{21} and S_{12} , Magnitude and Phase (cont)			Network analyzer with attenuators; f = frequency in GHz A_N = nominal attenuation in dB
Coaxial Type N (0.01 to 18) GHz $A_N \leq 20$ dB $(20 < A_N \leq 30)$ dB $(30 < A_N \leq 50)$ dB $(50 < A_N \leq 60)$ dB	0.060 dB 0.070 dB 0.20 dB 0.30 dB	$(0.090f + 0.60)$ ° $(0.090f + 0.60)$ ° $(0.090f + 0.60)$ ° $(0.090 f + 2.0)$ °	Attenuators within Agilent 85055A verification kit
Coaxial 3.5 mm (0.01 to 26.5) GHz $A_N \leq 20$ dB $(20 < A_N \leq 40)$ dB	$(0.0010 * f + 0.030)$ dB $(0.0010 * f + 0.050)$ dB	$(0.11f + 0.30)$ ° $(0.11 f + 0.30)$ °	Attenuators within Agilent 85053B verification kit
Coaxial 2.92 mm (0.01 to 40) GHz $A_N \leq 20$ dB $(20 < A_N \leq 50)$ dB	$(0.0016 * f + 0.030)$ dB $(0.0020 * f + 0.15)$ dB	$(0.11f + 0.50)$ ° $(0.11 f + 0.50)$ °	Attenuators within Anritsu model 3668 verification kit
Coaxial 2.4 mm (0.01 to 50) GHz $A_N \leq 20$ dB $(20 < A_N \leq 40)$ dB	$(0.0020 * f + 0.050)$ dB $(0.0012 * f + 0.080)$ dB	$(0.12f + 0.50)$ ° $(0.12 f + 0.50)$ °	Attenuators within Agilent 85057B verification kit

Features of note regarding the above word processing created version, include the following:

1. The above table still maintains a typical four column layout, yet the authors were forced into an almost complete redefinition of the typical columns, which would typically be:

Parameter/Equipment	Range	CMC[2] ($\pm$)	Comments

This is no cause for concern as long as the only intended use of the document is for it to be read by humans, but this kind of variation would present a great burden against developing a software based text parsing and data mining algorithm.

2. Note that there is a row/line alignment problem between the first column and the following columns that starts with the text "(0.01 to 26.5) GHz". To the left of this line the other columns should be blank; all of the data in the following columns should be shifted down by one row/line. This is a minor typographical error that is barely noticeable, but again, this kind of variation would present a great burden against developing a software based text parsing and data mining algorithm.

3. It is not clear if the frequency ranges listed in column one are merely comments on the typical frequency range of the corresponding connector type, or hard criteria.

4. Finally, it is not clear to what purpose is served by the mention of the particular attenuator sets in the last column.

What follows is a complete XML data file that fully adheres to the current version of the proposed' schema for the CMC section of an SOA document, which contains data intended to match as closely as possible to the above sample SOA table,   In comparison to the word processing version, the XML file is obviously much more verbose; extending for for nearly 7 pages and in a small font size. However, unlike the wordprocessing created version, every data element contained therein is fully defined to the extent required as to be completely unambiguous and fully comprehensible by data mining software.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<?xml-model href="http://schema.metrology.net/Uncertainty.sch" type="application/xml"
schematypens="http://purl.oclc.org/dsdl/schematron"?>
<unc:CMCs xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xmlns:xi="http://www.w3.org/2001/XInclude"
 xmlns:uom="http://schema.metrology.net/UnitsOfMeasure"
 xmlns:unc="http://schema.metrology.net/Uncertainty"
 xmlns:mml="http://www.w3.org/1998/Math/MathML"
xsi:schemaLocation="http://schema.metrology.net/Uncertainty.xsd">
  <unc:CMC>
    <unc:Category name="electrical:microwave"/>
    <unc:DUT_Type>Coaxial Microwave Attenuator</unc:DUT_Type>
    <unc:Process name="Measure S21 and S12 Magnitude and Phase"/>
    <unc:Technique name="Network Analyzer"/>
    <unc:Template>

      <unc:Selectors>
        <unc:Selector>
          <unc:Symbol>An</unc:Symbol>
          <unc:Quantity>attenuation</unc:Quantity>
          <unc:Description>Nominal Attenuation Value</unc:Description>
```

```xml
    </unc:Selector>
  </unc:Selectors>

  <unc:Formula heading="Magnitude">
    <unc:Function>A</unc:Function>
    <unc:SymbolDefinitions>
      <unc:Result>
        <unc:Quantity>attenuation</unc:Quantity>
      </unc:Result>
      <unc:Constant>
        <unc:Symbol>A</unc:Symbol>
        <unc:Quantity>attenuation</unc:Quantity>
      </unc:Constant>
    </unc:SymbolDefinitions>
  </unc:Formula>

  <unc:Formula heading="Phase">
    <unc:Function>m1*F + b1</unc:Function>
    <unc:SymbolDefinitions>
      <unc:Result>
        <unc:Quantity>phase-change</unc:Quantity>
      </unc:Result>
      <unc:Constant>
        <unc:Symbol>m1</unc:Symbol>
        <unc:Quantity>phase-change-per-unit-frequency</unc:Quantity>
        <unc:Description>Phase Frequency Coefficient</unc:Description>
      </unc:Constant>
      <unc:Constant>
        <unc:Symbol>b1</unc:Symbol>
        <unc:Quantity>phase-change</unc:Quantity>
        <unc:Description>Phase Uncertainty Floor</unc:Description>
      </unc:Constant>
      <unc:Input>
        <unc:Symbol>F</unc:Symbol>
        <unc:Quantity>frequency</unc:Quantity>
        <unc:Description>Frequency</unc:Description>
      </unc:Input>
    </unc:SymbolDefinitions>
  </unc:Formula>

  <unc:Conditions>
    <unc:Condition>
      <unc:Name>connector-type</unc:Name>
      <unc:Value>Coaxial Type N (0.01 to 18) GHz</unc:Value>
      <unc:Ranges>
        <unc:Selector>
          <unc:Symbol>An</unc:Symbol>

          <unc:Range>
            <unc:Start test="not applicable"/>
            <unc:End test="at">
              <unc:Value uom_alternative="decibels-power-ratio">20</unc:Value>
            </unc:End>
            <unc:Constant>
              <unc:Symbol>A</unc:Symbol>
              <unc:Value uom_alternative="decibels-power-ratio">0.060</unc:Value>
            </unc:Constant>
```

```xml
<unc:Constant>
   <unc:Symbol>m1</unc:Symbol>
   <unc:Value uom_alternative="degrees-per-gigahertz" >0.090</unc:Value>
</unc:Constant>
<unc:Constant>
   <unc:Symbol>b1</unc:Symbol>
   <unc:Value uom_alternative="degrees">0.60</unc:Value>
</unc:Constant>
</unc:Range>

<unc:Range>
   <unc:Start test="after">
      <unc:Value uom_alternative="decibels-power-ratio">20</unc:Value>
   </unc:Start>
   <unc:End test="at">
      <unc:Value uom_alternative="decibels-power-ratio">30</unc:Value>
   </unc:End>
   <unc:Constant>
      <unc:Symbol>A</unc:Symbol>
      <unc:Value uom_alternative="decibels-power-ratio">0.070</unc:Value>
   </unc:Constant>
   <unc:Constant>
      <unc:Symbol>m1</unc:Symbol>
      <unc:Value uom_alternative="degrees-per-gigahertz" >0.090</unc:Value>
   </unc:Constant>
   <unc:Constant>
      <unc:Symbol>b1</unc:Symbol>
      <unc:Value uom_alternative="degrees" >0.60</unc:Value>
   </unc:Constant>
</unc:Range>

<unc:Range>
   <unc:Start test="after">
      <unc:Value uom_alternative="decibels-power-ratio">30</unc:Value>
   </unc:Start>
   <unc:End test="at">
      <unc:Value uom_alternative="decibels-power-ratio">50</unc:Value>
   </unc:End>
   <unc:Constant>
      <unc:Symbol>A</unc:Symbol>
      <unc:Value uom_alternative="decibels-power-ratio">0.20</unc:Value>
   </unc:Constant>
   <unc:Constant>
      <unc:Symbol>m1</unc:Symbol>
      <unc:Value uom_alternative="degrees-per-gigahertz" >0.090</unc:Value>
   </unc:Constant>
   <unc:Constant>
      <unc:Symbol>b1</unc:Symbol>
      <unc:Value uom_alternative="degrees" >0.60</unc:Value>
   </unc:Constant>
</unc:Range>

<unc:Range>
   <unc:Start test="after">
      <unc:Value uom_alternative="decibels-power-ratio">50</unc:Value>
   </unc:Start>
   <unc:End test="at">
```

```xml
              <unc:Value uom_alternative="decibels-power-ratio">60</unc:Value>
            </unc:End>
            <unc:Constant>
              <unc:Symbol>A</unc:Symbol>
              <unc:Value uom_alternative="decibels-power-ratio">0.30</unc:Value>
            </unc:Constant>
            <unc:Constant>
              <unc:Symbol>m1</unc:Symbol>
              <unc:Value uom_alternative="degrees-per-gigahertz" >0.090</unc:Value>
            </unc:Constant>
            <unc:Constant>
              <unc:Symbol>b1</unc:Symbol>
              <unc:Value uom_alternative="degrees" >2.0</unc:Value>
            </unc:Constant>
          </unc:Range>

        </unc:Selector>
      </unc:Ranges>
    </unc:Condition>
  </unc:Conditions>

</unc:Template>

<unc:Template>

  <unc:Selectors>
    <unc:Selector>
      <unc:Symbol>An</unc:Symbol>
      <unc:Quantity>attenuation</unc:Quantity>
      <unc:Description>Nominal Attenuation Value</unc:Description>
    </unc:Selector>
  </unc:Selectors>

  <unc:Formula heading="Magnitude">
    <unc:Function>m2*F + b2</unc:Function>
    <unc:SymbolDefinitions>
      <unc:Result>
        <unc:Quantity>attenuation</unc:Quantity>
      </unc:Result>
      <unc:Constant>
        <unc:Symbol>m2</unc:Symbol>
        <unc:Quantity>attenuation-per-unit-frequency</unc:Quantity>
      </unc:Constant>
      <unc:Constant>
        <unc:Symbol>b2</unc:Symbol>
        <unc:Quantity>attenuation</unc:Quantity>
      </unc:Constant>
      <unc:Input>
        <unc:Symbol>F</unc:Symbol>
        <unc:Quantity>frequency</unc:Quantity>
      </unc:Input>
    </unc:SymbolDefinitions>
  </unc:Formula>

  <unc:Formula heading="Phase">
    <unc:Function>m1*F + b1</unc:Function>
    <unc:SymbolDefinitions>
```

```xml
      <unc:Result>
        <unc:Quantity>phase-change</unc:Quantity>
      </unc:Result>
      <unc:Constant>
        <unc:Symbol>m1</unc:Symbol>
        <unc:Quantity>phase-change-per-unit-frequency</unc:Quantity>
        <unc:Description>Phase Frequency Coefficient</unc:Description>
      </unc:Constant>
      <unc:Constant>
        <unc:Symbol>b1</unc:Symbol>
        <unc:Quantity>phase-change</unc:Quantity>
        <unc:Description>Phase Uncertainty Floor</unc:Description>
      </unc:Constant>
      <unc:Input>
        <unc:Symbol>F</unc:Symbol>
        <unc:Quantity>frequency</unc:Quantity>
        <unc:Description>Frequency</unc:Description>
      </unc:Input>
    </unc:SymbolDefinitions>
</unc:Formula>

<unc:Conditions>

  <unc:Condition>
    <unc:Name>connector-type</unc:Name>
    <unc:Value>Coaxial 3.5 mm (0.01 to 26.5) GHz</unc:Value>
    <unc:Ranges>
      <unc:Selector>
        <unc:Symbol>An</unc:Symbol>

        <unc:Range>
          <unc:Start test="not applicable"/>
          <unc:End test="at">
            <unc:Value uom_alternative="decibels-power-ratio">20</unc:Value>
          </unc:End>
          <unc:Constant>
            <unc:Symbol>m1</unc:Symbol>
            <unc:Value uom_alternative="degrees-per-gigahertz">0.11</unc:Value>
          </unc:Constant>
          <unc:Constant>
            <unc:Symbol>b1</unc:Symbol>
            <unc:Value uom_alternative="degrees">0.30</unc:Value>
          </unc:Constant>
          <unc:Constant>
            <unc:Symbol>m2</unc:Symbol>
            <unc:Value uom_alternative="decibels-power-ratio-per-gigahertz">0.0010</unc:Value>
          </unc:Constant>
          <unc:Constant>
            <unc:Symbol>b2</unc:Symbol>
            <unc:Value uom_alternative="decibels-power-ratio">0.030</unc:Value>
          </unc:Constant>
        </unc:Range>

        <unc:Range>
          <unc:Start test="after">
            <unc:Value uom_alternative="decibels-power-ratio">20</unc:Value>
          </unc:Start>
```

```xml
        <unc:End test="at">
          <unc:Value uom_alternative="decibels-power-ratio">40</unc:Value>
        </unc:End>
        <unc:Constant>
          <unc:Symbol>m1</unc:Symbol>
          <unc:Value uom_alternative="degrees-per-gigahertz">0.11</unc:Value>
        </unc:Constant>
        <unc:Constant>
          <unc:Symbol>b1</unc:Symbol>
          <unc:Value uom_alternative="degrees">0.30</unc:Value>
        </unc:Constant>
        <unc:Constant>
          <unc:Symbol>m2</unc:Symbol>
          <unc:Value uom_alternative="decibels-power-ratio-per-gigahertz">0.0010</unc:Value>
        </unc:Constant>
        <unc:Constant>
          <unc:Symbol>b2</unc:Symbol>
          <unc:Value uom_alternative="decibels-power-ratio">0.050</unc:Value>
        </unc:Constant>
      </unc:Range>

    </unc:Selector>
  </unc:Ranges>
</unc:Condition>

<unc:Condition>
  <unc:Name>connector-type</unc:Name>
  <unc:Value>Coaxial 2.92 mm (0.01 to 40) GHz</unc:Value>
  <unc:Ranges>
    <unc:Selector>
      <unc:Symbol>An</unc:Symbol>

      <unc:Range>
        <unc:Start test="not applicable"/>
        <unc:End test="at">
          <unc:Value uom_alternative="decibels-power-ratio">20</unc:Value>
        </unc:End>
        <unc:Constant>
          <unc:Symbol>m1</unc:Symbol>
          <unc:Value uom_alternative="degrees-per-gigahertz">0.11</unc:Value>
        </unc:Constant>
        <unc:Constant>
          <unc:Symbol>b1</unc:Symbol>
          <unc:Value uom_alternative="degrees">0.50</unc:Value>
        </unc:Constant>
        <unc:Constant>
          <unc:Symbol>m2</unc:Symbol>
          <unc:Value uom_alternative="decibels-power-ratio-per-gigahertz">0.0016</unc:Value>
        </unc:Constant>
        <unc:Constant>
          <unc:Symbol>b2</unc:Symbol>
          <unc:Value uom_alternative="decibels-power-ratio">0.030</unc:Value>
        </unc:Constant>
      </unc:Range>

      <unc:Range>
        <unc:Start test="after">
```

```
        <unc:Value uom_alternative="decibels-power-ratio">20</unc:Value>
      </unc:Start>
      <unc:End test="at">
        <unc:Value uom_alternative="decibels-power-ratio">50</unc:Value>
      </unc:End>
      <unc:Constant>
        <unc:Symbol>m1</unc:Symbol>
        <unc:Value uom_alternative="degrees-per-gigahertz">0.11</unc:Value>
      </unc:Constant>
      <unc:Constant>
        <unc:Symbol>b1</unc:Symbol>
        <unc:Value uom_alternative="degrees">0.50</unc:Value>
      </unc:Constant>
      <unc:Constant>
        <unc:Symbol>m2</unc:Symbol>
        <unc:Value uom_alternative="decibels-power-ratio-per-gigahertz">0.0020</unc:Value>
      </unc:Constant>
      <unc:Constant>
        <unc:Symbol>b2</unc:Symbol>
        <unc:Value uom_alternative="decibels-power-ratio">0.15</unc:Value>
      </unc:Constant>
    </unc:Range>

  </unc:Selector>
 </unc:Ranges>
</unc:Condition>

<unc:Condition>
  <unc:Name>connector-type</unc:Name>
  <unc:Value>Coaxial 2.4 mm (0.01 to 50) GHz</unc:Value>
  <unc:Ranges>
    <unc:Selector>
      <unc:Symbol>An</unc:Symbol>

      <unc:Range>
        <unc:Start test="not applicable"/>
        <unc:End test="at">
          <unc:Value uom_alternative="decibels-power-ratio">20</unc:Value>
        </unc:End>
        <unc:Constant>
          <unc:Symbol>m1</unc:Symbol>
          <unc:Value uom_alternative="degrees-per-gigahertz">0.12</unc:Value>
        </unc:Constant>
        <unc:Constant>
          <unc:Symbol>b1</unc:Symbol>
          <unc:Value uom_alternative="degrees">0.50</unc:Value>
        </unc:Constant>
        <unc:Constant>
          <unc:Symbol>m2</unc:Symbol>
          <unc:Value uom_alternative="decibels-power-ratio-per-gigahertz">0.0020</unc:Value>
        </unc:Constant>
        <unc:Constant>
          <unc:Symbol>b2</unc:Symbol>
          <unc:Value uom_alternative="decibels-power-ratio">0.050</unc:Value>
        </unc:Constant>
      </unc:Range>
```

```xml
            <unc:Range>
              <unc:Start test="after">
                <unc:Value uom_alternative="decibels-power-ratio">20</unc:Value>
              </unc:Start>
              <unc:End test="at">
                <unc:Value uom_alternative="decibels-power-ratio">40</unc:Value>
              </unc:End>
              <unc:Constant>
                <unc:Symbol>m1</unc:Symbol>
                <unc:Value uom_alternative="degrees-per-gigahertz">0.12</unc:Value>
              </unc:Constant>
              <unc:Constant>
                <unc:Symbol>b1</unc:Symbol>
                <unc:Value uom_alternative="degrees">0.50</unc:Value>
              </unc:Constant>
              <unc:Constant>
                <unc:Symbol>m2</unc:Symbol>
                <unc:Value uom_alternative="decibels-power-ratio-per-gigahertz">0.0012</unc:Value>
              </unc:Constant>
              <unc:Constant>
                <unc:Symbol>b2</unc:Symbol>
                <unc:Value uom_alternative="decibels-power-ratio">0.08</unc:Value>
              </unc:Constant>
            </unc:Range>

          </unc:Selector>
        </unc:Ranges>
      </unc:Condition>

    </unc:Conditions>

  </unc:Template>

</unc:CMC>
<xi:include href="http://testsite2.callabsolutions.com/UnitsOfMeasure/UOM_Database.xml"/>
</unc:CMCs>
```