

BioCTS 2013

BioCTS for AN-2011

NIEM XML Encoded Transactions User Guide

**NIST/ITL CSD Biometric Conformance Test Software for
ANSI/NIST-ITL 2011 NIEM XML Encoded Transactions**

July 2013

Dylan Yaga

NIST/ITL CSD Lead Software Designer

Fernando Podio

NIST/ITL CSD Project Manager

Christofer J. McGinnis

ID Technology Partners (NIST Associate) Software Developer

National Institute of Standards and Technology (NIST)

Information Technology Laboratory (ITL)

Computer Security Division (CSD)

Contents

1. Disclaimer.....	3
2. Overview	4
2.1. Requirements.....	4
3. Relevant Standards, Standards Requirements and Additional Test Assertions	5
3.1. Relevant Standards and Specifications	5
3.2. Additions to the Standard Requirements	5
4. Testing Phases.....	8
5. Required Changes to the ANSI/NIST-ITL 1-2011 Schema	9
6. Guide	9
6.1. Download and Installation	9
6.2. Running the Conformance Test Architecture	10
6.3. Conformance Test Architecture Features.....	11
6.3.1. New Features	11
6.3.2. Batch Testing.....	13
6.3.3. Editor.....	16
6.3.4. Options.....	17
7. Anatomy of a Result.....	19
8. References	21
9. Annex A.....	22

1. Disclaimer

NIST/ITL BioCTS

For AN-2011 NIEM XML Encoded Transactions

October 2010

The software was developed by the National Institute of Standards and Technology (NIST), an agency of the Federal Government. Pursuant to Title 15 United States Code Section 105, works of NIST are not subject to copyright protection in the United States and are considered to be in the public domain. Thus, the software may be freely reproduced and used. Please explicitly acknowledge the National Institute of Standards and Technology as the source of the software.

This software is released by NIST as a service and is expressly provided "AS IS." NIST MAKES NO WARRANTY OF ANY KIND, EXPRESS, IMPLIED OR STATUTORY, INCLUDING, WITHOUT LIMITATION, THE IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, NON-INFRINGEMENT AND DATA ACCURACY. NIST DOES NOT REPRESENT OR WARRANT THAT THE OPERATION OF THE SOFTWARE WILL BE UNINTERRUPTED OR ERROR-FREE, OR THAT ANY DEFECTS WILL BE CORRECTED.

NIST does not warrant or make any representations regarding the use of the software or the results thereof, including but not limited to the correctness, accuracy, reliability or usefulness of the software. By using this software or by incorporating this software into another product, you agree to hold harmless the United States Government for any and all damages or liabilities that arise out of such use.

Certain trade names and company products are mentioned in the text or identified. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products are necessarily the best available for the purpose. With the exception of material marked as copyrighted, information presented in this document is considered public information and may be distributed or copied. Use of appropriate byline/photo/image credits is requested.

2. Overview

This document describes the features of biometric conformance test software developed by NIST/ITL Computer Security Division (CSD). BioCTS for ANSI/NIST-ITL 1-2011 (AN-2011) National Information Exchange Model (NIEM)-conformant XML Encoded Transactions (BioCTS for AN-2011 XML) implements all test assertions listed as NIEM-conformant in NIST Special Publication (SP) 500-295 *Conformance Testing Methodology for ANSI/NIST-ITL 1-2011, Data Format for the Interchange of Fingerprint, Facial & Other Biometric Information* [1] in addition to new and updated assertions which were developed after the publication of NIST SP 500-295. This publication includes tables of requirements and test assertions (over 1,200 test assertions necessary to addressing the test requirements specified in the ANSI/NIST-ITL 1-2011 standard). A revision of NIST SP 500-295 is planned reflecting the required additional and updated test assertions that were found to be necessary during the development of the test tool.

2.1. Requirements

- Supported Microsoft[®] Operating Systems:
 - Windows[®] XP[™] Service Pack 3
 - Windows[®] Vista[™] Service Pack 2
 - Windows[®] 7[™] Service Pack 1
- Microsoft[®] .NET 4.0 Framework:
 - Microsoft[®] .NET 4.0 Web Installer (<http://www.microsoft.com/en-us/download/details.aspx?id=17851>)*
 - Microsoft[®] .NET 4.0 Stand Alone Installer (<http://www.microsoft.com/en-us/download/details.aspx?id=17718>)*

(*) Links working as of 7/16/2013

3. Relevant Standards, Standards Requirements and Additional Test Assertions

3.1. Relevant Standards and Specifications

The relevant standard is ANSI/NIST-ITL 1-2011, NIST Special Publication 500-290 *Data Format for the Interchange of Fingerprint, Facial & Other Biometric Information*. Only the requirements that apply to NIEM-conformant transactions, known as an Exchange Package, are relevant to BioCTS for AN-2011 XML.

3.2. Additions to the Standard Requirements

The test tool implements all test assertions listed as NIEM-conformant in NIST SP 500-295, as well as the additional and updated test assertions found required after the publication of NIST SP 500-295. Test Limitations with respect to the Standard

There are requirements specified in the AN-2011 standard that do not have test assertions and are not implemented within the software. The Table below identifies and provides justification for all exceptions..

Exception	Section	AN-2011 Requirement Summary	Justification
Domain Names / Application Profile Specifications	5.3.2	Data contained in this record shall conform in format and content to the specifications of the domain name(s) as listed in Field 1.013 Domain name / DOM found in the Type-1 record, if that field is in the transaction. The default domain is NORAM. Field 1.016 Application profile specifications / APS allows the user to indicate conformance to multiple specifications. If Field 1.016 is specified, the Type-2 record must conform to each of the application profiles. A DOM or APS reference uniquely identifies data contents and formats. Each domain and application profile shall have a point of contact responsible for maintaining this list. The contact shall serve as a registrar and maintain a repository including documentation for all of its common and user-specific Type-2 data fields. As additional fields are required by specific agencies for their own applications, new fields and definitions may be registered and reserved to have a specific meaning. When this occurs, the domain or application profile registrar is responsible for registering a single definition for each number used by different members of the domain or application profile.	The format and content of the record are defined by the DOM or APS. Each DOM and APS has related record-content definitions that may be updated. The evolving nature of the DOM and APS definitions and nature of using registrars means that the requirements are not defined in the base standard. ¹
	6	An implementation domain, coded in Field 1.013 Domain name / DOM of a Type-1 record as an optional field, is a group of agencies or organizations that have agreed to use preassigned data fields with specific meanings (typically in Record Type-2) for	Since the “transaction must conform to each profile” included in the field, and those profiles are defined by the listed

¹ Requirements related to user, profile, or domain-specific information are not within the scope of conformance testing to the base AN-2011 standard, and therefore are not tested by the BioCTS for AN-2011 software developed to test AN-2011 Traditional and NIEM XML Encoded Transactions..

		<p>exchanging information unique to their installations. The implementation domain is usually understood to be the primary application profile of the standard.</p> <p>New to this version of the standard, Field 1.016 Application profile specifications / APS allows multiple application profiles to be referenced. The organization responsible for the profile, the profile name and its version are all mandatory for each application profile specified. A transaction must conform to each profile that is included in this field. It is possible to use Field 1.016 and / or Field 1.013.</p> <p>A specified implementation domain and specified application profiles must all have the same definition for fields, subfields and information items that are contained in the transaction.</p>	<p>agency, the CTS would have to retrieve the latest requirements from the agency.¹</p>
Alternate Character Sets	5.6, Table 2	<p>Field 1.015 Character encoding/DCS is an optional field that allows the user to specify an alternate character encoding... Field 1.015 Character encoding/DCS contains three information items: the character encoding set index/ CSI, the character encoding set name/CSN, and the character encoding set version/CSV. The first two items are selected from the appropriate columns of Table 2.</p>	<p>Table 2 lists ASCII, UTF-16, UTF-8, and UTF-32 as possible encodings. However, the table also allows “User-defined” character encoding sets.¹</p>
Alternate Coordinate System	7.7.3, Table 4	<p>The ninth information item is the geodetic datum code / GDC10. It is an alphanumeric value of 3 to 6 characters in length. This information item is used to indicate which coordinate system was used to represent the values in information items 2 through 7. If no entry is made in this information item, then the basis for the values entered in the first eight information items shall be WGS84, the code for the <i>World Geodetic Survey 1984 version - WGS 84 (G873)</i>. See Table 4 for values.</p>	<p>Table 4 lists 22 coordinate systems and the option to include “Other” types as well. It is not feasible for the CTS to test conformance to all coordinate systems, specifically those that are listed by the user under “Other”.¹ The CTS tests for conformance to WGS84 because it is the default coordinate system used in the base standard.</p>
	7.7.3	<p>A fourteenth optional information item geographic coordinate other system identifier / OSI allows for other coordinate systems. This information items specifies the system identifier. It is up to 10 characters in length. Examples are:</p> <ul style="list-style-type: none"> • MGRS (Military Grid Reference System) • USNG (United States National Grid) • GARS (Global Area Reference System) • GEOREF (World Geographic Reference) • LANDMARK (e.g. hydrant) and position relative to the landmark. <p>A fifteenth optional information item, is the geographic coordinate other system value / OCV. It shall only be present if OSI is present in the record. It can be up to 126 characters in length. If OSI is LANDMARK, OCV is free text and may be up to 126 characters. For details on the formatting of OCV for the other coordinate systems shown in OSI as examples, see http://earth-</p>	<p>While some examples of other coordinate systems are listed in the standard (MGRS, USNG, GARS, GEOREF, LANDMARK), those values are not all-inclusive, and the user may indicate other coordinate systems that are not listed¹.</p>

		info.nga.mil/GandG/coordsys/grids/referencesys.html	
Subject Acquisition Profiles SAP/FAP/IAP	7.7.5, Table 8, Table 9, Table 10	A subject acquisition profile is used to describe a set of characteristics concerning the capture of the biometric sample. These profiles have mnemonics SAP for face, FAP for fingerprints and IAP for iris records.	It is not feasible to test if the image was captured under the conditions specified by the SAP, FAP or IAP level as defined in Tables 10 through 13. However, the fields will be tested for valid level values.
Open and Closed Paths	7.8	<p>Several Record Types define open paths (also called contours or polylines) and / or closed paths (polygons) on an image. They are comprised of a set of vertices. For each, the order of the vertices shall be in their consecutive order along the length of the path, either clockwise or counterclockwise. (A straight line of only two points may start at either end). A path may not have any sides crossing. No two vertices shall occupy the same position. There may be up to 99 vertices.</p> <p>An open path is a series of connected line segments that do not close or overlap. A closed path (polygon) completes a circuit. The closed path side defined by the last vertex and the first vertex shall complete the polygon. A polygon shall have at least 3 vertices.</p> <p>The contours in Record Type-17: Iris image record can be a circle or ellipse. A circle only requires 2 points to define it (See Table 16).</p> <p>There are two different approaches to the paths in this standard. The 2007 and 2008 version of the standard used paths for Field 14.025: Alternate finger segment position(s) / ASEG. That approach has been retained in this version for all paths except in the Extended Feature Set (EFS) of Record Type-9. The EFS adopted an approach expressing the path in a single information item, which is different than that used in other record types.</p>	<p>Further research is needed to determine the feasibility of testing for:</p> <ul style="list-style-type: none"> -simple, plane figure -no sides crossing -no interior holes

4. Testing Phases

The BioCTS for AN-2011 NIEM XML conformance test suite (CTS) design was a result of careful analysis of the base requirements in the ANSI/NIST-ITL 1-2011 standard, the NIEM-Encoding specific requirements, and the NIEM XML Schema definitions. The NIEM schema is able to verify that the Exchange Package data is formatted and ordered correctly, but it does not ensure conformance to all of the base requirements. This observation is reflected in the BioCTS for AN-2011 NIEM XML design, which validates against the schema, reports the validation as a single result, and continues testing all of the required test assertions.

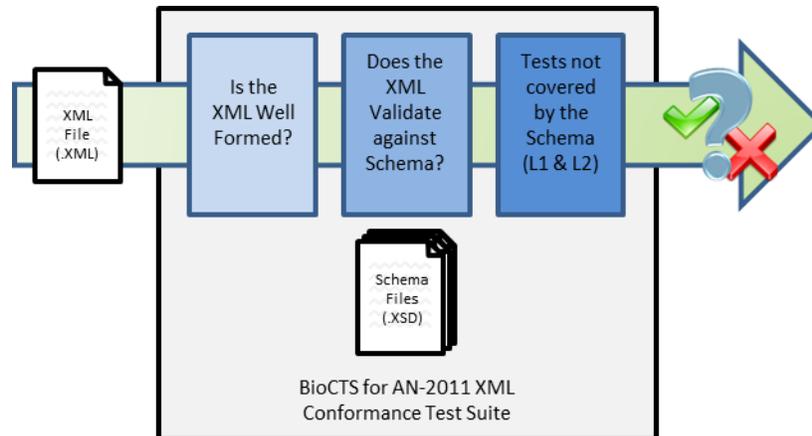


Figure 1 - High Level CTS Program Flow

The CTS has three test phases:

1. **Determination of whether the XML file is Well-Formed.**

An XML file is Well-Formed when it is syntactically correct, and follows the rules of XML documents [2] [3]. Without a Well-Formed XML file, further testing would yield potentially unusable results.

*If the XML file fails this phase of testing, the CTS **will not continue** to the next phase.*

2. **Validating the XML file against the specified schema file.**

BioCTS for AN-2011 XML will attempt to validate the XML file against the specified schema file (BioCTS allows for a user-defined schema file, but also provides a default), and will report as many errors as possible.

*The CTS, regardless of whether the XML file passes or fails this phase of testing, **will continue** to the next phase.*

3. **Assertion testing for AN-2011 base requirements.**

The final phase is testing against the base requirements of ANSI/NIST-ITL 1-2011 that are not covered by the XML Schema file validation. These tests include, but are not limited to:

- a. Valid Value tests
- b. Character Count tests
- c. Relationship tests between XML elements
- d. Basic Image Validation

After all three phases of testing is complete for the XML file, the test results are aggregated and an overall result (Pass or Fail) is determined. For the overall result of an XML file to be reported as “Pass” there must be no result of “Error” or “Critical” in any of the XML file’s results.

5. Required Changes to the ANSI/NIST-ITL 1-2011 Schema

The Section C.5.1 in the ANSI/NIST-ITL 1-2011 standard specifies:

To the extent possible, the schema define data types and constraints that enforce the allowable content rules of the base standard. Nevertheless, the XML schema may not strictly enforce the allowable content. The base standard defines allowable content, and its requirements shall be met by implementers regardless of encoding method.

After a careful analysis of the base requirements of the ANSI/NIST-ITL 1-2011 standard and the Schema files provided [4], several discrepancies were found. As a result of these discrepancies it was concluded that the tool could not use the posted schema; the modified schema files that were developed concurrently with the test tool are available at the time of this writing at the test tool release web site.

The discrepancies between the schema provided and the base ANSI/NIST-ITL 1-2011 requirements can be categorized as:

1. The schema is **LESS** restrictive than the base requirements, and may allow additional values other than what is defined in the base requirements.
 - *In this case, the schema does not prevent the base requirement from passing.*
 - Additional tests have been implemented in the BioCTS for AN-2011 XML software to ensure that the values found within Exchange Package conform to the base requirements.
2. The schema is **MORE** restrictive than the base requirements, and may not allow certain values that are allowed in the base requirements.
 - *In this case, the schema does prevent the base requirement from passing.*
 - Additions and modifications were made to the schema files in the cases where the schema **prevented** the base requirements from passing. The BioCTS team modified the schema files distributed with the AN-2011 standard. These modifications are reflected in the schema file included with the BioCTS for AN-2011 XML software.
 - A summary of the schema modifications can be found in Annex A.

6. Guide

6.1. Download and Installation

Download the installer from the website

http://www.nist.gov/itl/csd/biometrics/biocta_download.cfm .

After the download completes, run the install program:

BioCTS_AN2011_1.1.XXXX.YYYYY.exe

(XXXX.YYYYY are the specific version numbers for the release) and follow the on-screen instructions presented in the dialog boxes.

6.2. Running the Conformance Test Architecture

To run the CTA software from the Start Menu:

Select **All Programs** then select **NIST BioCTS**, then select **AN-2011** and click on **BioCTS_AN2011**.

6.3. Conformance Test Architecture Features

6.3.1. New Features

Two new features have been added to BioCTS for AN-2011 (Traditional and NIEM XML Encoded Transactions) since beta version 1.0.4771.15880 (released January 2013) Text Log Output Search and Basic Statistics of the Test Results.

6.3.1.1. Text Log Output Search

Searching through an entire Text Log can be difficult, especially when a Transaction has multiple Records within it. To help alleviate this problem, without forcing the user to open an external tool, BioCTS for AN-2011 has implemented an internal search feature. To launch the search window, click the magnifying glass icon or use the keyboard command CTRL+F.

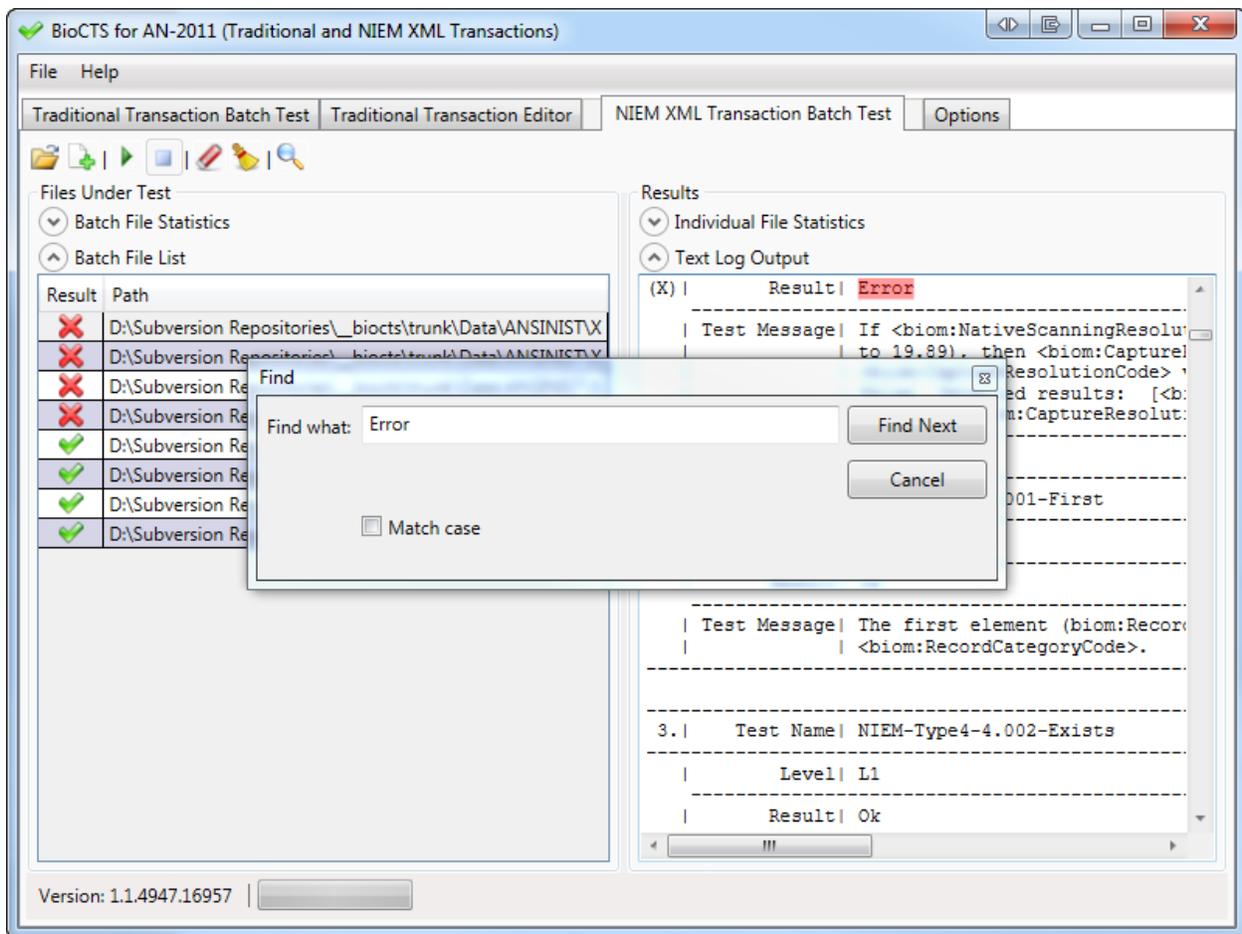


Figure 2 - BioCTS Text Log Output Search Feature

6.3.1.2. Statistics

In previous releases of BioCTS it was difficult to quantify many test metrics such as the number of files loaded and tested, the number of tests performed, and how many tests resulted in an error or warning. Basic statistics are now available including:

- **Selected File Statistics:** Details about how many tests were performed and a breakdown of what types of results were found. When a file is selected within the Batch XML Test tab, the statistics are updated to reflect that file.
- **Overall Batch Statistics:** Details about how many files were loaded into the Batch File tester, how many files were tested, and how many files were considered “Passing” or “Failing”.

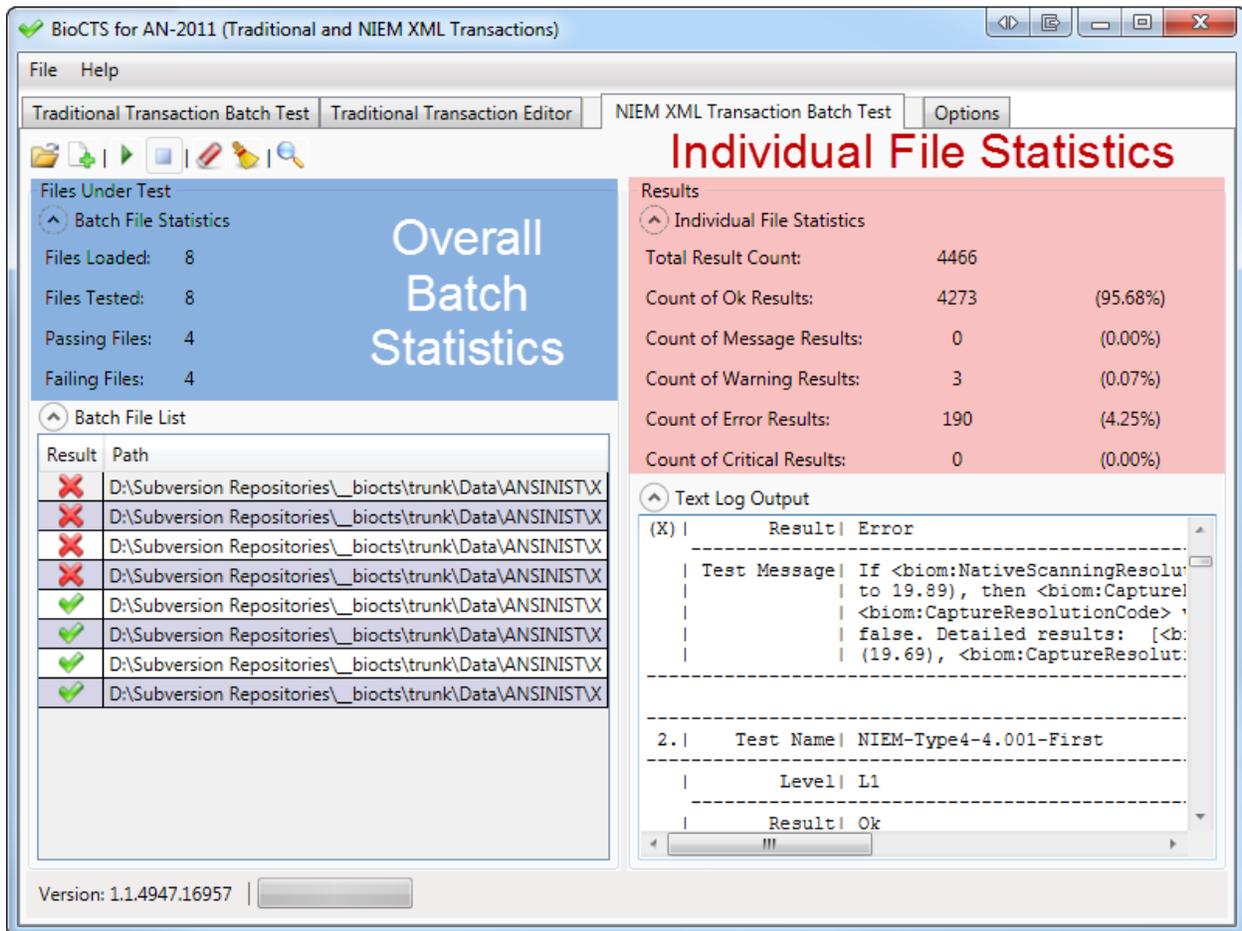


Figure 3 - BioCTS Highlighting the Statistic Summaries

6.3.2. Batch Testing

The “Batch Test” tab allows multiple transactions (files) to be tested in groups, and displays the overall results for each transaction in the “Files Under Test” pane.

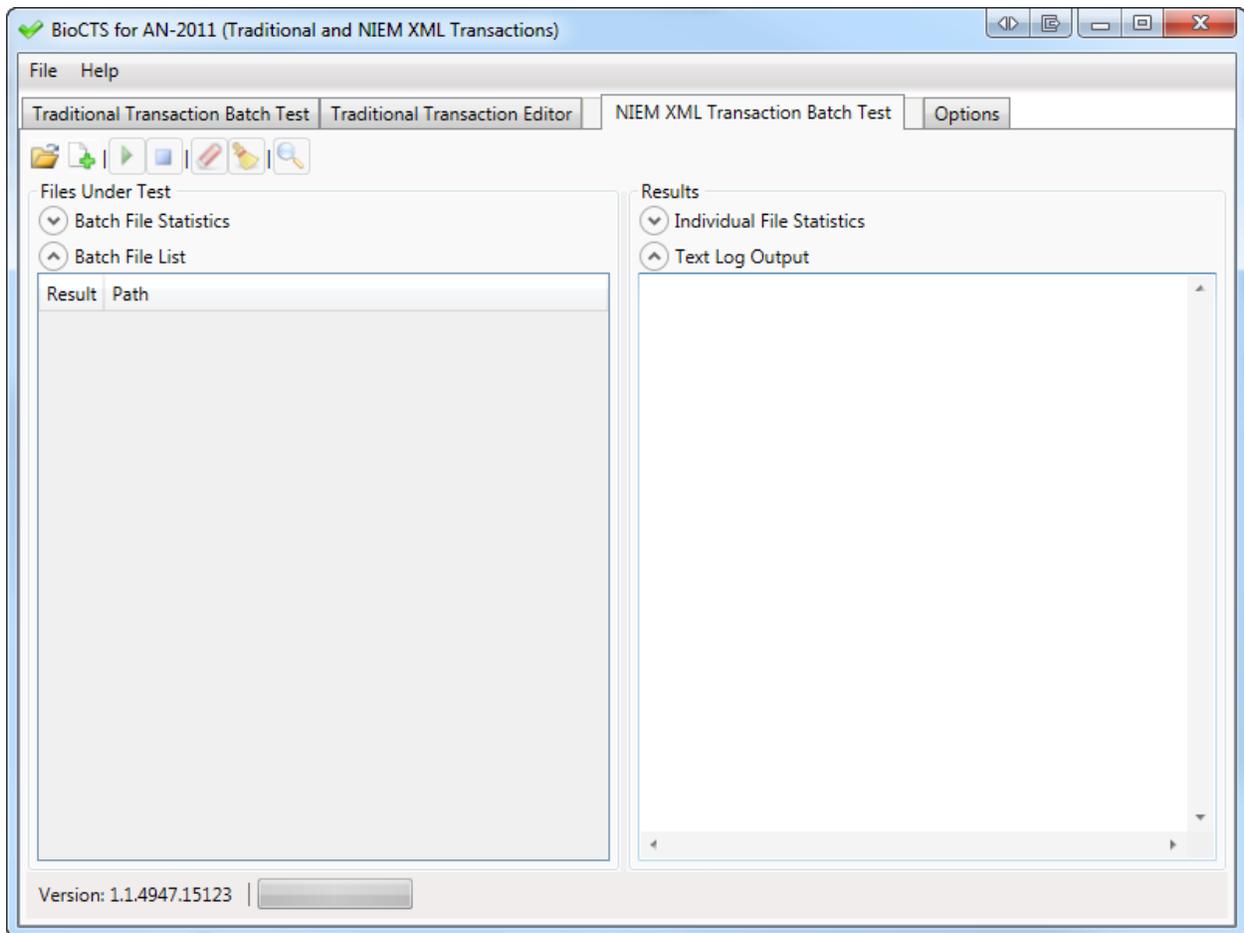


Figure 4 - BioCTS Initial XML Batch Test Tab

Several files can be loaded at once.

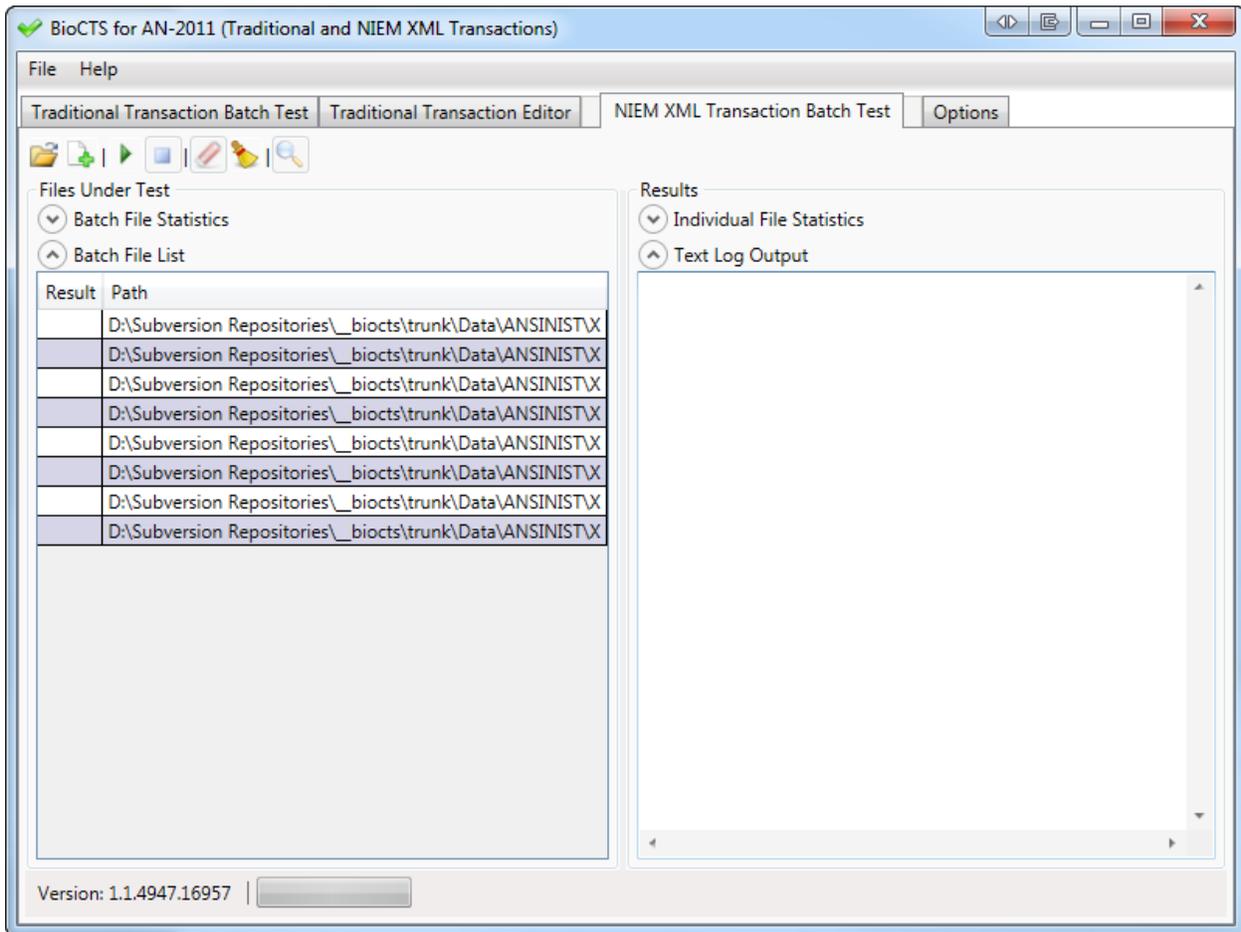


Figure 5 - BioCTS with several files loaded

The “Batch Test” tab will display the transaction’s overall result with either:

- ❌ - Overall Result of Fail
- ✅ - Overall Result of Pass

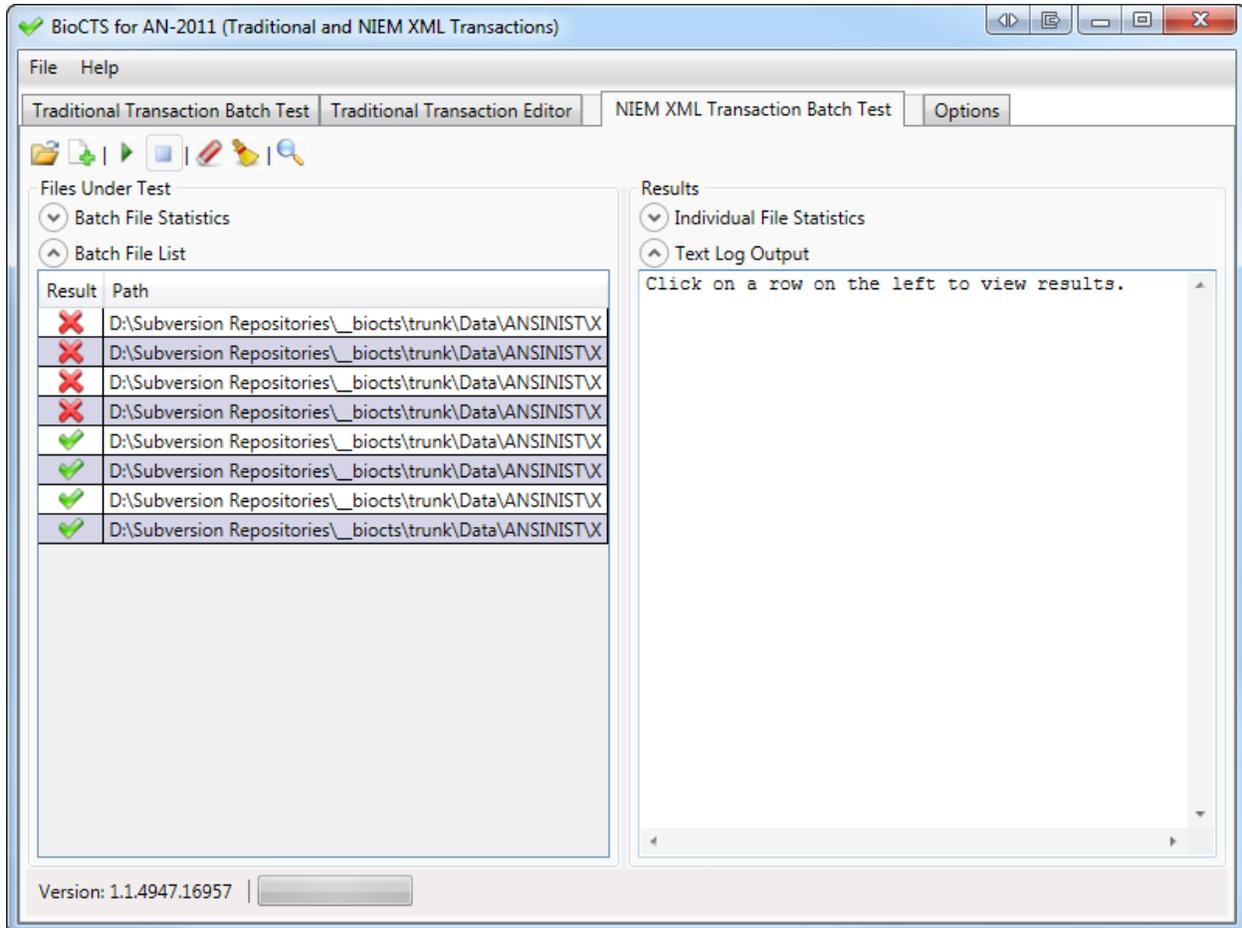


Figure 6 - BioCTS with several files loaded, displaying their overall results

Textual output results for each transaction can be viewed by clicking on the desired filename in the “File Under Test” pane. The complete textual results are displayed in the pane to the right.

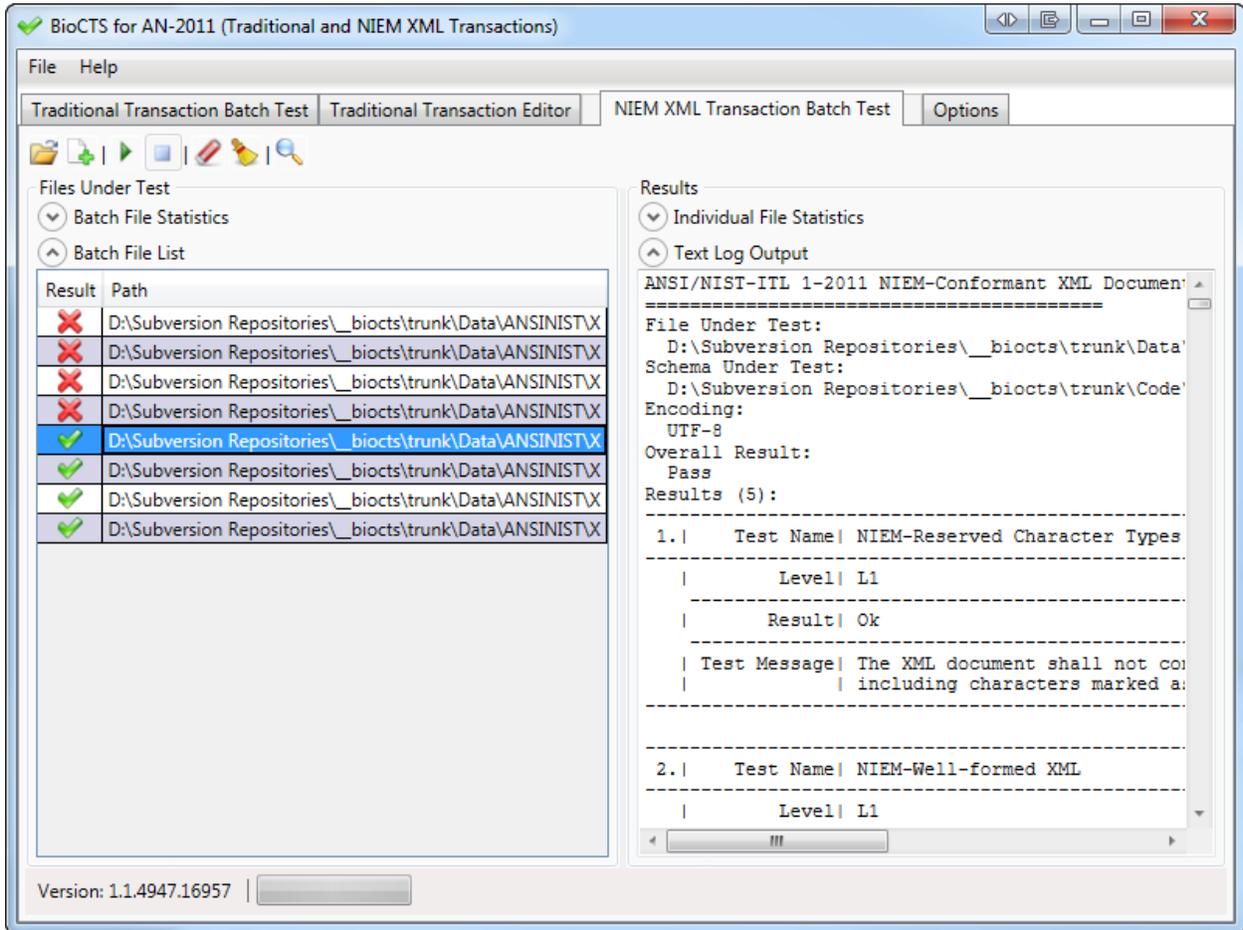


Figure 7 - BioCTS displaying the Text Log Output of a file that has been tested

6.3.3. Editor

There are many tools (both freely and commercially available) that allow for editing of the Text-based XML files. Currently, BioCTS for AN-2011 XML Encoded Transactions does not include an editor specifically for the XML Transactions.

6.3.4. Options

The “Options” tab provides options for output file type, constraint schema path and Text and XML Log Output save location.

BioCTS for AN-2011 is capable of two types of output; first is the Text Log Output, which is always generated during batch testing (for Traditional and XML). The second is an optional XML Output, which will generate an XML Output log that includes the same information found in the Text Log, but in a machine-readable format.

Please note that because of the amount of details provided in the Log Outputs, the size of these logs can be large.

Output is generated in a Time-stamped folder located in the output directory. The time-stamped folder format:

- yyyy – 4 digit year (e.g. 2012)
- MM – 2 digit month (e.g. 10)
- dd – 2 digit day (e.g. 31)
- HH – 2 digit hour in 24-hour scale (e.g. 13)
- mm – 2 digit minutes (e.g. 59)
- ss – 2 digit seconds (e.g. 22)

Examples:

- Text Output will be generated in the directory:
C:\Users\dyaga\Desktop\BioCTS for AN-2011 Output\2012.10.31.13.59.22\Text Output
- XML Output will be generated in the directory:
C:\Users\dyaga\Desktop\BioCTS for AN-2011 Output\2012.10.31.13.59.22\XML Output

In addition to the options above, there is an option to load in a User-Defined Constraint Schema for the XML Encoded Transaction Testing (as allowed in section C.3 of the ANSI/NIST-ITL 1-2011 standard). Clicking the “Default” button will restore the original schema file provided with BioCTS for AN-2011 XML.

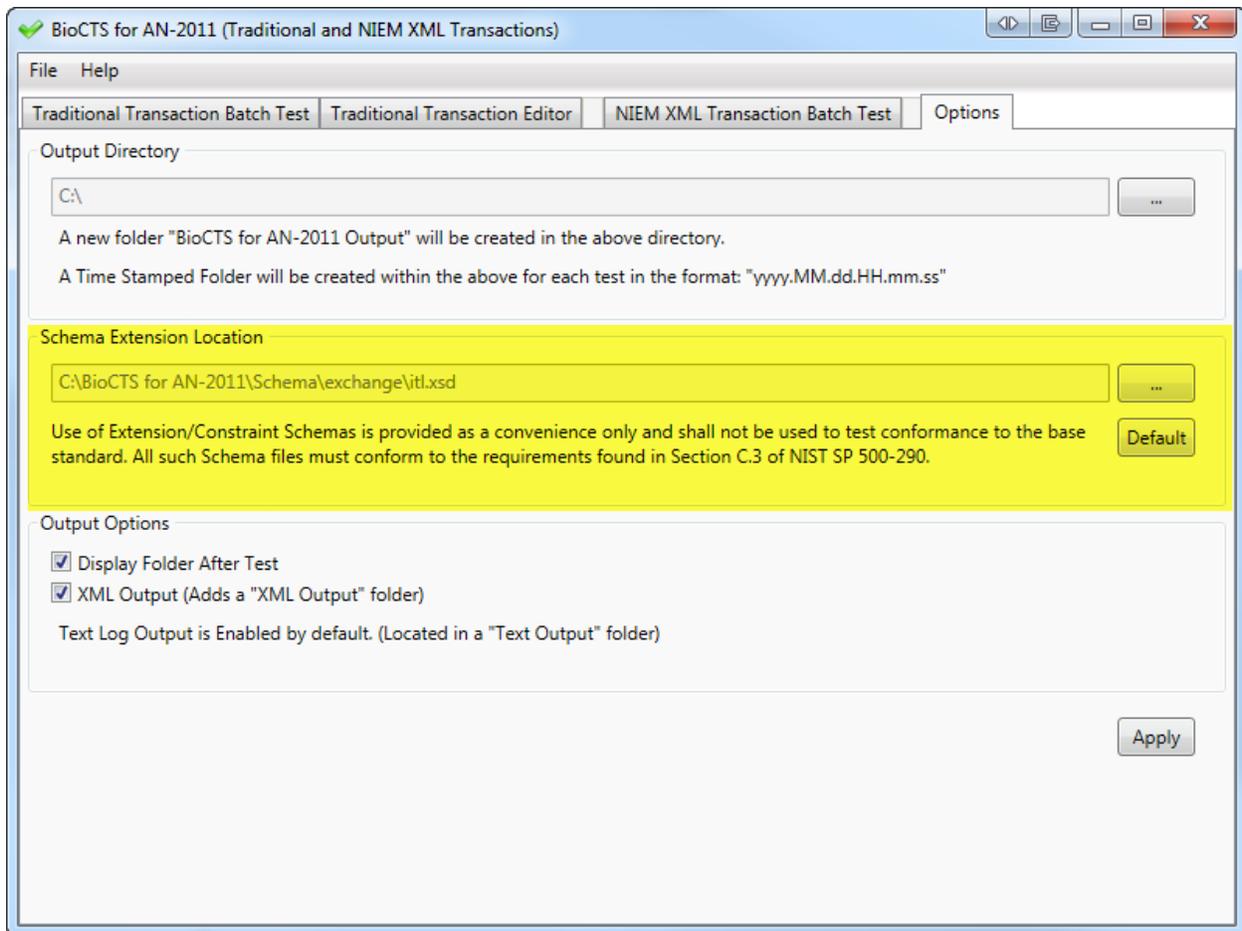


Figure 8 - The Options Tab with Schema Selection Highlighted

7. Anatomy of a Result

Every test performed in BioCTS for AN-2011 generates a Result – passing or failing. In an effort to enhance the readability and clarity of the Results, the textual output formatting has been modified since the initial release of BioCTS for AN-2011.

The Text Log Output Result formatting is:

```
-----  
1. |   Test Name | NIEM-Type1-1.001-First  
-----  
   |           | Level | L1  
-----  
   |           | Result | Ok  
-----  
   | Test Message | Is the first element <biom:RecordCategoryCode>?  
-----
```

Figure 9 - A Text Log Result

The XML Log Output Result format:

```
<Result>  
  <Level>L1</Level>  
  <Message>The first element (biom:RecordCategoryCode) in the record shall be <biom:RecordCategoryCode>.</Message>  
  <Results>Ok</Results>  
  <Test>NIEM-Type1-1.001-First</Test>  
</Result>
```

Figure 10 - An XML Log Result

A Result consists of, and clearly identifies:

- Test Name – The Test Names match up with what is found in NISTIR 500-295
- Test Level
 - Parse – A parse level test deals with any test that relates to parsing of the information; tests with this level happen before the data is inspected during L1, L2 and L3 testing.
 - L1 – A Level 1 test, which tests for values, lengths, and character counts of the data.
 - L2 – A Level 2 test, which tests for relationships between Fields, Subfields, Information Items and Records.
 - L3 – A Level 3 test, which tests to see if the data specified is consistent with the biometric sample presented.
- Test Result
 - Ok – The test was unable to find an error. This does not necessarily mean that this portion of data was without error; just that the tests could not find error.
 - Message – The test was unable to find an error; however the test found it necessary to convey an additional message.
 - Warning – The test was unable to find an error; however there may be an aspect that warrants further investigation.
 - Error – The test was able to find an error.
 - Critical – The test was able to find an error; this error was critical enough that it may impede further testing.

- Test Message – The message can contain any additional information to clarify the test.

8. References

1. NIST Special Publication 500-295 "Conformance Testing Methodology for ANSI/NIST-ITL 1-2011, Data Format for the Interchange of Fingerprint, Facial & Other Biometric Information" (Release 1.0), August 2012, Editors: *Fernando L. Podio, Dylan Yaga, Christofer J. McGinnis*
2. "Well-formed document" *Wikipedia: The Free Encyclopedia*. Wikimedia Foundation, Inc., 23 April 2013. Web. 30 May 2013.
<http://en.wikipedia.org/wiki/Well-formed_document>
3. "Extensible Markup Language (XML) 1.0" World Wide Web Consortium (W3C), 10 February 1998. Web. 30 May 2013
<<http://www.w3.org/TR/1998/REC-xml-19980210#dt-wellformed>>
4. ANSI/NIST-ITL 1-2011 Sample Data Files, including schemas, sample traditional and XML transactions. <http://biometrics.nist.gov/cs_links/standard/ansi_2011/ANSi-Files.zip>

9. Annex A

Below are the changes that have been made to the schema files as of the time of writing.

Additions and Modifications to: `subset\niem\domains\biometrics\1.0\biometrics.xsd`

- **Added:** Values (128 to 999) to the Simple Type `CharacterSetIndexCodeSimpleType`
Reason: The values listed in the original schema prevented the use of “User-defined character encoding sets” as specified in Table 4 of the AN-2011 standard which are specified as values 128 to 999.
- **Added:** A `minOccurs="0"` to the element `biom:FaceImageAcquisitionProfile` of the complex type `FaceImageType`.
Reason: As specified in Table 57 of the AN-2011 standard, Field 10.013 has a min occurrence of 0, max occurrence of 1. By not having the `minOccurs="0"` the schema was requiring this element to always be present.
- **Added:** The value 18 for “Unknown friction ridge” to the Simple Type `FingerPositionCodeSimpleType`.
Reason: Value 18 is specified as a valid value in Table 8 of the AN-2011 standard, but was not present in the schema.
- **Added:** A new simple type: `AlphabeticStringSimpleType` which allows the alphabetic characters and space (characters `[a-zA-Z\s]`), for zero or more times.
Reason: It was used as a basis for the following type:
 - **Added:** A new Type `OneToSixteenCharacterAlphabeticStringSimpleType` – This puts a min length of 1 and max length of 16 to the `AlphabeticStringSimpleType`
Reason: This new type was used to amend the `TransactionCategoryCode` which specifies a list of enumerated values. Just having the list of values was preventing the base requirement of having a user-defined field of Alphabetic strings 1 to 16 characters long (Field 1.004 in Table 22 of the AN-2011 standard)
- **Modified:** The Element `FingerprintImageStitchedIndicator` to only accept the character “Y”.
Reason: The Element was defined in the schema as a Boolean type (`true` or `false`). This prevented the base requirement of only allowing a single alphabetic character of “Y” (Field 14.027 in Table 71 of the AN-2011 standard).