Using Automation for S1000D Conversion and QA

Naveh Greenberg,
Director, U.S. Defense Development,
Data Conversion Laboratory
Agenda

• Is Fully Automated Solution Possible?
• Planning
• Project Start-up Methodology
• The process for achieving automation
• Q&A
Is Fully Automated Solution Possible?

• Depending on who you ask....
• Depending on what you consider is good output....
• WYSIWYG?
• Is it more reliable than hand-tagging?
Planning

“If I had eight hours to chop down a tree, I'd spend six sharpening my ax.”

• Things you must know:
  • Who are the stakeholders? Who is the real client (they have to take part in the process)?
  • Budget.
  • Schedule (or at least know major milestones & completion date).
  • CSDB, IETM, sustainment tools.

• Things you need:
  • Legacy data, business rules, schemas, samples.
  • The “right people”. It’s a team work.

• Things to do:
  • QA plan
  • Agree on rules.
  • Have samples before moving forward.
  • Pilot and/or Limited Production Run.
"If I had eight hours to chop down a tree, I'd spend six sharpening my ax." – Abraham Lincoln
Workflow With Automated Conversion

Step 1: Document Submission
Step 2: Document Logging
Step 3: DMRL Creation & Preparation

Step 4: Automated Conversion
Step 5: Parsing
Step 6: Editorial Review

Step 7: Final Quality Assurance (QA)
Step 8: Integration of XML and CGM into an IETM
Step 9: Delivery

Step 10: Feedback and Resubmission

Automation can be applied in every step of the process (not limited to step 4).
Automation During Analysis – If Legacy in ASCII

- Finding exact or similar text will help you when mapping to Data Modules

  
  A17-1 2010 U002 025 line: 2054
  engineering test: a test carried out by or witnessed by a registered or licensed professional engineer, testing laboratory, or certifying organization to ensure conformance to Code requirements.

  <Exact matches>
  A17-1 CSA B44 Edn 2013 000505 U100 2013 015 line: 2164
  A17-1 Handbook Edn 2010 032114 000501 u004 2010 033 line: 19514
  A17-3 Edn 2011 000502 U001 2011 045 line: 7406

  253P63 line: 90
  This is the safety alert symbol. When you see this symbol on your pump or in this manual, look for one of the following signal words and be alert to the potential for personal injury.
  This is the safety {safety-alert} alert symbol. When you see this symbol on your pump or in this manual, look for one of the following signal words and be alert to the potential for personal injury. {injury: injury!}
  This is the safety alert symbol. When you see this symbol on your pump or in this manual, look for one of the following signal words and be alert to the potential for personal injury {injury: injury!}
  S502 line: 69
  S506 line: 56
  S509 line: 54

- It will also help to detect applicability and inconsistencies

  
  AG 1b U004 054 line: 14401
  FD-5220 OTHER WELDS Fillet welds, butt welds, and seal welds shall be visually inspected in accordance with AA-6331.
  FD-5220 (FH-5220) OTHER WELDS Fillet welds, butt welds, and seal welds shall be visually inspected in accordance with AA-6331.
  FD-5220-FH-5220 OTHER WELDS Fillet welds, butt welds, and seal welds shall be visually inspected in accordance with AA-6331.
  AG 1b U004 054 line: 22357
Proofreading – Extracted Data

Sample Document Text

Sample Page

DETAILED DESCRIPTION OF THE INVENTION

The invention is directed to compounds according to formula (I):

where R1 is the group –XYZ or
X is phenyl, heteroaryl, 1,2,3,4-tetrahydronaphthalenyl, or 2,3-dihydro-1H-indenyl,
**Proofreading – Visual & Automated**

**C-1440**

Sealing 0-Ring

<table>
<thead>
<tr>
<th>String</th>
<th>Char</th>
<th>Count</th>
<th>Line #</th>
</tr>
</thead>
<tbody>
<tr>
<td>115-Vac</td>
<td>#1</td>
<td>1</td>
<td>117</td>
</tr>
<tr>
<td>125psi</td>
<td>#1</td>
<td>2</td>
<td>554,788</td>
</tr>
<tr>
<td>1070T10</td>
<td>#0</td>
<td>3</td>
<td>3049,3437</td>
</tr>
<tr>
<td>1070T10-7</td>
<td>#0</td>
<td>1</td>
<td>3083</td>
</tr>
<tr>
<td>1070TO-5</td>
<td>#0</td>
<td>1</td>
<td>3095</td>
</tr>
<tr>
<td>0-ring</td>
<td>#0</td>
<td>1</td>
<td>7434</td>
</tr>
<tr>
<td>4OF76Y56-72014</td>
<td>#0</td>
<td>1</td>
<td>8179</td>
</tr>
</tbody>
</table>

---

**VARIANTS**

Not symmetrical space

<table>
<thead>
<tr>
<th>String</th>
<th>Char</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>A01 -HR</td>
<td>#1</td>
<td>1</td>
</tr>
<tr>
<td>1 -Reference</td>
<td>#1</td>
<td>1</td>
</tr>
<tr>
<td>1- 223645</td>
<td>#1</td>
<td>1</td>
</tr>
</tbody>
</table>

---

**Space before punctuation mark**

<table>
<thead>
<tr>
<th>String</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>removed.</td>
<td>#2</td>
</tr>
<tr>
<td>21 .</td>
<td>#2</td>
</tr>
<tr>
<td>of .</td>
<td>#2</td>
</tr>
</tbody>
</table>
Preparing for conversion

2-3. GENERAL

A permanent record of the services, repairs required, and how to complete them. (Army) See DA Pam 734.

2-4. CLEANING INSTRUCTIONS

a. Cleaning is an after-operation service that requires the application of cleaning methods and procedures. (Army)
2-1. COMMON TOOLS AND EQUIPMENT. For authorized common tools and equipment refer to Modified Table of Organization and Equipment (†) applicable to your unit, Navy, Air Force and Coa Guard users are authorized to use the following common tools:

<table>
<thead>
<tr>
<th>Task</th>
<th>Paragraph</th>
<th>Tool</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maintenance of 9mm Pistol</td>
<td>2-11</td>
<td>None required</td>
</tr>
<tr>
<td>Maintenance of Receiver Assembly</td>
<td>2-12</td>
<td>Screwdriver</td>
</tr>
<tr>
<td>Maintenance of Magazine Catch Assembly</td>
<td>2-13</td>
<td>Screwdriver</td>
</tr>
</tbody>
</table>

2-2. SPECIAL TOOLS AND SUPPORT EQUIPMENT. There are no special tools for this item. Tools and test equipment are listed in appendix B of this manual. There is no Test, Measurement, and Diagnostic Equipment (TMDE) for this item.

2-3. REPAIR PARTS. Repair parts are listed and illustrated in appendix C of this manual.

2-13. MAINTENANCE OF MAGAZINE CATCH ASSEMBLY.

This task covers removal, reversing and installation.

INITIAL SETUP

Tools and Special Tools

<table>
<thead>
<tr>
<th>Equipment Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pistol grips removed</td>
</tr>
</tbody>
</table>

Tool Set, Small Arms (SC 5180-95-CL-A07)

WARNING

Make sure weapon is clear and there are no obstructions in the barrel or chamber.

REMOVAL

Remove the magazine catch assembly (1) by pushing in and to the rear with the fingertip, on the side opposite the magazine catch assembly button (2). The magazine catch assembly will then drop out.

NOTE

To reverse the magazine catch, install the button on the op

DMRL will map to DM but, when possible, automation is needed to restructure data
Collect Titles to a “Pre analysis Spreadsheet” or a DMRL

Frame Maker:

Chapter 1: XML
Working with XML files

About XML

XML is a language designed to mark content so it can be output in multiple ways. Like an HTML file, an XML file includes tags that surround and organize content. However, unlike HTML, XML tags do not sp

SGML/XML:

<chapter><title>
<section><title>
<para0><title>
<proc><title>

WORD:

General

A. This procedure has these tasks:

(1) TASK 31-31-11-000-802, Flight Data Recorder
(2) TASK 31-31-11-400-802, Flight Data Recorder

B. [SF] The flight dta recorder is installed between STA compartment.

TASK 31-31-11-000-802

Fligh Data Recorder Removal

A. References

PDF:

1. General

A. This procedure has these tasks:

(1) TASK 31-31-11-000-802, Flight Data Recorder
(2) TASK 31-31-11-400-802, Flight Data Recorder

B. [SF] The flight dta recorder is installed between STA compartment.

TASK 31-31-11-000-802

Flight Data Recorder Removal

A. References
<table>
<thead>
<tr>
<th>Para No</th>
<th>Original Title</th>
<th>DMC</th>
<th>Type</th>
<th>Comment</th>
<th>Tools</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOREWORD</td>
<td>SAMPLESYSTEM-A-00-00-00A-018A-D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>INTRODUCTION.</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>MANUAL ORGANIZATION.</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHAP 3</td>
<td>PREPARATION FOR USE AND RESHIPMENT</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SECT I</td>
<td>PREPARATION FOR USE</td>
<td>SAMPLESYSTEM-A-00-00-00A-006A-A</td>
<td>Desc</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>INTRODUCTION.</td>
<td>------</td>
<td></td>
<td>Repeat: 3.2, 3.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1.1</td>
<td>Nomenclature Setup and Initialization Instructions</td>
<td>SAMPLESYSTEM-A-00-00-00A-007A-A</td>
<td>proc</td>
<td></td>
<td>Wrench[1];Hammer[1]</td>
<td>Para 3.3;Chapter 2; Para 3.1.1, TO 14-77-001</td>
</tr>
<tr>
<td>SECT II</td>
<td>PREPARATION FOR RESHIPMENT</td>
<td>SAMPLESYSTEM-A-00-00-00A-530A-A</td>
<td>proc</td>
<td></td>
<td>Screw[4];Washer[4]; O-ring[2]</td>
<td></td>
</tr>
<tr>
<td>3.2</td>
<td>COMMON TOOLS.</td>
<td>------</td>
<td></td>
<td>Reconst</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.3</td>
<td>CONDITIONS AND METHODS OF RESHIPMENT.</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHAP 8</td>
<td>ILLUSTRATED PARTS BREAKDOWN</td>
<td>SAMPLESYSTEM-A-00-00-00A-018B-D</td>
<td>Desc</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SECT I</td>
<td>INTRODUCTION</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.1</td>
<td>INTRODUCTION.</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.2</td>
<td>MAINTENANCE PARTS LIST.</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SECT II</td>
<td>MAINTENANCE PARTS LIST</td>
<td>SAMPLESYSTEM-A-00-00-00A-941A-D</td>
<td>IPD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SECT III</td>
<td>NUMERICAL INDEX</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SECT IV</td>
<td>REFERENCE DESIGNATOR INDEX</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GLOSSARY</td>
<td>SAMPLESYSTEM-A-00-00-00A-006A-D</td>
<td>Desc</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Automated Conversion Software

Regardless of the legacy you’re starting with, and the target XML schema you’re going to, some items are handled exactly the same.

Maintain legacy footprint until the last step of the process.
Reporting, Audit & Reconciliation

Invocation: airqc(Revision1.2.0) -s -pgraphics -mk:\P27078\sw\qc\qc#.ctl WP002.sgm
* ERR * EOF, tags not closed: wptablelist wplep wpilluslist , msg=906
* ERR * Table orient=port width is not standard: 7 (not 6.5) Inches; line #:106, msg=337
* ERR * Table orient=port width is not standard: 7 (not 6.5) Inches; line #:196, msg=337
* WARN * <emphasis> ends on space: 'Help '; line #:164, msg=569
* WARN * Check space: '/ ' in <para>When the symbol &lt;graphic1&gt; is placed between a procedure/step
* WARN * Check title text 'FOREWORD' in <title>FOREWORD</title>; line #:13, msg=121
* WARN * Check if not WP ref 'See WP 007' in <para>See WP 007 (Maintenance Procedures for Operators)
* WARN * Graphic with boardno='703_0356' occur 2 times; line #:526, msg=252
* WARN * Graphic with boardno='703_0418' occur 2 times; line #:3443, msg=252
* WARN * Table in body/rear without title text: <table frame="all">; line #:91, msg=320
* WARN * Table with width less then 3.5 inches spans page Table width 2.32 < 3.50000138888889 Inches; :
* WARN * Check if not figure or figure ref 'FIG 3/' in <entry>FIG 3/></entry>; line #:77, msg=509
* WARN * Check if not reference 'to paragraph 3.203).' in <para>button. The system saves the cable swi
* WARN * Content is absent ' ' in <title></title>; line #:92, msg=604
* WARN * Element has no content ' ' in <title></title>; line #:92, msg=606
* WARN * Step without label: Wait until Windows safely shuts down and the screen goes black. 'label'; :
* WARN * Content is absent ' ' in <para></para>; line #:75, msg=704
* INFO * **************** ** QC CTL file: k:\P27078\sw\qc\qcLEG.ctl ****************, msg=800
Reporting, Audit & Reconciliation

errs/2 Attribute @dmcode is not allowed on element <dmCode> (See http://www.w3.org/TR/xmlschema-1/#cvc-complex-type clause 4)
errs/2 Attribute @inschvl is not allowed on element <levelledPara> (See http://www.w3.org/TR/xmlschema-1/#cvc-complex-type clause 4)
errs/3 Attribute @label is not allowed on element <listItem> (See http://www.w3.org/TR/xmlschema-1/#cvc-complex-type clause 4)
errs/2 Attribute @leftInd is not allowed on element <tgroup> (See http://www.w3.org/TR/xmlschema-1/#cvc-complex-type clause 4)
errs/5 FORG0001: The value "" is not a valid NCName (See http://www.w3.org/TR/xmlschema-1/#cvc-complex-type clause 4)
errs/3 FORG0001: Value "unknown" contravenes the enumeration facet "modify, delete, add" of the type changeTypeAtt
errs/1 In content of element <action>: The content model does not allow element <brk> to appear here. Expected one or more elements of type changeTypeAtt
errs/3 In content of element <action>: The content model does not allow element <note> to appear here. Expected one or more elements of type changeTypeAtt
errs/7 In content of element <action>: The content model does not allow element <note> to appear here. Expected one or more elements of type changeTypeAtt
errs/1 In content of element <action>: The content model does not allow element <query> to appear here. Expected one or more elements of type changeTypeAtt
errs/1 In content of element <dmIdent>: The content model does not allow element <dmcode> to appear here. Expected one or more elements of type changeTypeAtt
errs/1 In content of element <dmIdent>: The content model does not allow element <dmcode> to appear here. Expected one or more elements of type changeTypeAtt
errs/11 In content of element <isolationProcedureEnd>: The content model does not allow element <dmRef> to appear here. Expected one or more elements of type changeTypeAtt
errs/5 In content of element <levelledPara>: The content model does not allow element <sequentialList> to appear here. Expected one or more elements of type changeTypeAtt
errs/11 In content of element <levelledPara>: The content model does not allow element <brk> to appear here. Expected one or more elements of type changeTypeAtt
errs/8 In content of element <para>: The content model does not allow element <brk> to appear here. Expected one or more elements of type changeTypeAtt
errs/4 In content of element <para>: The content model does not allow element <brk> to appear here. Expected one or more elements of type changeTypeAtt
errs/1 In content of element <para>: The content model does not allow element <note> to appear here. Expected one or more elements of type changeTypeAtt
errs/2 In content of element <para>: The content model does not allow element <note> to appear here. Expected one or more elements of type changeTypeAtt
errs/4 In content of element <para>: The content model does not allow element <table> to appear here. Expected one or more elements of type changeTypeAtt
errs/3 In content of element <para>: The content model does not allow element <table> to appear here. Expected one or more elements of type changeTypeAtt
errs/1 In content of element <para>: The content model does not allow element <table> to appear here. Expected one or more elements of type changeTypeAtt
errs/1 In content of element <para>: The content model does not allow element <table> to appear here. Expected one or more elements of type changeTypeAtt
errs/2 In content of element <proceduralStep>: The content model does not allow element <caution> to appear here. Expected one or more elements of type changeTypeAtt
errs/9 In content of element <proceduralStep>: The content model does not allow element <caution> to appear here. Expected one or more elements of type changeTypeAtt
errs/4 In content of element <proceduralStep>: The content model does not allow element <figure> to appear here. Expected one or more elements of type changeTypeAtt
errs/2 In content of element <proceduralStep>: The content model does not allow element <note> to appear here. Expected one or more elements of type changeTypeAtt
errs/1 In content of element <proceduralStep>: The content model does not allow element <warning> to appear here. Expected one or more elements of type changeTypeAtt
errs/8 In content of element <proceduralStep>: The content model does not allow element <proceduralStep> to appear here. Expected one or more elements of type changeTypeAtt
errs/2 Required attribute @assYCode is missing (See http://www.w3.org/TR/xmlschema-1/#cvc-complex-type clause 4)
errs/2 Required attribute @disasYCode is missing (See http://www.w3.org/TR/xmlschema-1/#cvc-complex-type clause 4)
errs/2 Required attribute @disasYCodeVariant is missing (See http://www.w3.org/TR/xmlschema-1/#cvc-complex-type clause 4)
errs/2 Required attribute @infoCode is missing (See http://www.w3.org/TR/xmlschema-1/#cvc-complex-type clause 4)
errs/2 Required attribute @infoCodeVariant is missing (See http://www.w3.org/TR/xmlschema-1/#cvc-complex-type clause 4)
errs/2 Required attribute @itemLocationCode is missing (See http://www.w3.org/TR/xmlschema-1/#cvc-complex-type clause 4)
errs/2 Required attribute @modelIdentCode is missing (See http://www.w3.org/TR/xmlschema-1/#cvc-complex-type clause 4)
errs/2 Required attribute @subSubSystemCode is missing (See http://www.w3.org/TR/xmlschema-1/#cvc-complex-type clause 4)
Q&A

Naveh Greenberg
Director, U.S. Defense Development,
Data Conversion Laboratory
(718) 307-5758
ngreenberg@dclab.com
@dclaboratory