



Converting Legacy ATA 100 Interleaf Manuals to S1000D/ATA 1000BR: A User Experience

Normand Montour



SUMMARY

- 01 Business context – Why convert ?
- 02 Current situation
- 03 What we found in the source documents
- 04 Key ATA BREX rules that guided our decisions
- 05 Conversion highlights

Introduction

This presentation will discuss an actual user experience of converting legacy ATA 100 manuals written in Interleaf to an S1000D format using the Spec 1000BR as the BREX. We will describe the challenges encountered while converting the file format, focusing on specific technical issues and solutions.

- **Remaining conformant to the BREX rules, including decisions on whether to adhere strictly to BREX or to choose documented exceptions.** And what drove to these decisions.
- **How to adapt a legacy effectivity model to a corresponding applicability model,** even if it was decided not to proceed.
- **Retaining effectivity information while conforming to the S1000D content model.**
- **Illustrations conversion to consider,** from bitmap and vectorial PDF formats to conformant CGM, and technical challenges involved.
- **How to efficiently infer and transform "paper links"** from the original Interleaf documents into valid S1000D link markup, both internal and external.
- **Publishing of the massive corpus of DMs, illustrations, and Publishing Modules,** using the formatting engine and stylesheet provided by the CSDB vendor.

01 - Business context – Why convert ?

Legacy documents in ATA « print » format authored in Interleaf

- Expensive licences and maintenance support
- Difficulty in finding adequately trained personnel to update documents
- Multiple scattered portions of documents in various support storage

Legal / delivery requirements

- End client demands S1000D content
- End clients demands ATA BREX conformance

Why we got involved?

- Expertise in Technical Writing, ATA, S1000D, translation, illustration
- We have the licences, the skills, and the knowledge
- The client can get rid of their Interleaf system altogether

02 - Current situation

Thousands of legacy source documents in various states

- Incoherent classification
- Difficult to open
- Often illustrations are stored separately (when they are not lost)
- Created by various releases of the software
- Sometimes only a PDF rendition is available

Content characteristics

- Designed exclusively for a paper (PDF) ATA rendition
- Allowed structures incompatible with S1000D (for ex. tables within figures)

Time constraint

- The client wants to have the conversion done in a narrow calendar
- Need to discontinue licences (\$\$) asap

03 - What we found in the source documents

Content characteristics

- “Paper links”, not marked up in any way in the source
- ATA effectivity statements in various places (DM-level, inline, in table columns, in page footer (sometimes as a graphic) without any specific effectivity markup (paper only)
- Revision marks throughout the publication
- Inconsistent naming of ID attributes (sometimes non-conformant) for textual sections
- The WDM had no source material (interleaf), and everything had to be taken from a pdf
- Many of the interleaf files were missing, and the only source were scanned pages added to the pdf
- Inconsistent formatting between manual types
- Many of the tables had to be changed for figures and then have updated internal references.
- Issues on if links within graphics are considered valid S1000D

03 - What we found in the source documents (continued)

Format consideration

- Only Interleaf-ASCII or HTML output could be used for conversion
(we used the HTML output for simplicity and easier inference search)
- Tables withing figures will be converted to Legend (if two columns) or moved after the figure if more than 2 columns.

BREX consideration

- Invalid infoCodes found (were used to mark chapter numbers)
- Unusual structures (procedures running deep ; ≥ 8 or 9 levels
- AIPC discrepancies

04 - Key ATA BREX rules that guided our decisions

ATA BREX Constraints

- Does not allow “incomplete” linking (@targetTitle attribute and <internalRef> element cannot contain textual content (or spaces).)
- Does not allow “incomplete” applicability (<applicCrossRefTableRef> element is required within the <pmStatus> and <dmStatus> element.)
- Strict rule on some content (e.g. depth of procedures ≤5)
- Strict naming of infoCode (some invalid)
- Context-sensitive content (ATA BREX does not allow for enumerated list within caution and warnings)

Decisions made

- Infer linking based on paper references in the manuals and in the AIPC
- Infer both internal (fig , table, proc) and external (Chapter, DMs, etc.) linking
- Procedures imbedded greater than level 5 transformed into nested sequential lists
- Inference of Cautions and Warnings to be move to appropriate CIRs
- Parts callouts we not tagged as hotspots to synchronize between illustrated Parts data/figures

04 - Key ATA BREX rules that guided our decisions (continued)

BREX accepted Exceptions (to be tested)

- Effectivity statements to be retained but not converted to applicability.
- All revision marks will be ignored (Issue 0 of new pub)
- The BREX default manual type "CMM" will be overridden by the client's manual type "AMM", thus :
manualType will be "AMM"
(`<externalPubCode pubCodingScheme="manualType">AMM</externalPubCode>`) not "CMM" as the default dictates
- Did not use CIR for Parts Vendor Info – It was deemed too complicated, even if the quality of the manual would have been improved.

05 - Conversion highlights

Lessons learned

- Converting from HTML (akin to plain text) was adequate.
We could have had better results converting from Interleaf ASCII, but the learning curve was considered impractical.
- Coherent nomenclature of ID attributes (DM, Proc, Table , Figure, CIR entry, etc.) is key when trying to infer links from textual (« paper ») references
- Link inference is based on pattern recognition to be reconciled with IDs referential (database)
- Some character-based errors in parsing (for ex. N-dash vs. Hyphen, zero for « o », etc.)
- Use of extensive Schema/BREX checker is essential
- Use of CIRs (by inference) improve consistency of publication (we found a few typos that way)
- Converting PDF graphics to vectorial (CGM) can yield visual fidelity, but not necessarily updatable drawings
- Retaining ATA-style effectivity as plain textual content without conversion to an applicability model may lead to confusion and make further updates difficult

