Document Integrity and Assurance

Michael Wood
Group 1 Software
What is Document Integrity

Definition

in·teg·ri·ty (n-tgr-t)n.
1. Steadfast adherence to a strict moral or ethical code.
2. The state of being unimpaired; soundness.
3. The quality or condition of being whole or undivided; completeness.

The goal of performing Document Integrity is to ensure that the document creation and change processes creates completely sound, correct, and valid documents.
The Importance of Document Integrity

- Documents are a significant percentage of business communications with customers.
- Businesses are recognizing an increased importance in effective communication – but errors in documents destroys their effectiveness, as well as eroding a Corporation’s image.
The Value of Document Integrity

- Testing represents up to 60% of document development time (per: Madison Advisors)

**Improving this results in:**

- Affects time-to-market for new products
- Paper, printing, and shipping costs for test copy
- Process should catch and repair document errors
- Reduced errors means less customer confusion, reducing customer service call volume and expense, and increases customer satisfaction
How can Testing be 60% of Time?

Iterative Development Cycle

Testing is in the Critical Path, and is a Blocking operation.
Mini-Poll: How Do YOU Perform Document Integrity Today?

- Production-Printed Copy Proofing
- PDF Copy Proofing
- Print Stream Compare Software
- Other
- None?
Issue with Existing Validation Methods

Manual Task

Forces compromises between Thoroughness and Speed

Intrinsically risky: The human eye can “gloss over” certain differences easily. An 8 looks like a 0. A 0 looks like an O.

Boring work introduces even more error
How Do you Validate Documents?

Compare a “Known Good” versus “Trial”
Assuming the documents are supposed to be identical

Compare using “Light Box” approach

✓ Identifies movement issues
✓ Identifies new or missing elements
How Do you Validate Documents?

**Compare a “Known Good” versus “Trial”**

If there are known, expected differences between the copies

1. Make a “Check List” of things to check
2. Check each item for each document
3. Repeat until finished

✓ Can be as thorough as needed.
✓ Can be as complex as needed.
How Do you Validate Documents?

Extract Information and Compare to Production Databases

1. Extract Information from each Document, or save to a Log / Journal File
2. Compare Field Data to Database
3. Repeat until finished

✓ Can handle different data layouts – for example changing order of columns in the output, or changing the font size
✓ Good for a complete “closed loop” production system
How Do you Validate Documents?

Extract Information and Compare to Extracted Information from Good Copy

1. Extract Information from each Document in Trial Copy, or save to a Log / Journal File
2. Repeat for Good Copy
3. Compare Records to Each Other.

✓ Can validate computations that take place as part of the composition process
When is Document Integrity needed?

Migration To…

• Another business system
• Another composition system
• An upgraded composition system version
  • Minor version upgrade…
  • Major version upgrade…
• Another print format
• New Host platform (i.e. Mainframe - Client Server)
When is Document Integrity needed?

Business Drivers...

- Redesigned document
- New product, service offering
- New pricing structure, discount plans
- Implementation of Colour
- Implementing Transpromo
- Regulatory requirements
- Change to remit address / PO Box
- M&A
Introducing...

StreamSure Print Stream Validation System
StreamSure Features…

Robust Comparison Engine

• Compare entire print streams
• Conditional comparisons – can compare certain components of the document only when a pre-condition is satisfied, just like humans do
• Doesn’t have to compare every element – since not everything is always supposed to be the same
• Can compare AFP to Metacode, Line Data
StreamSure Features…

Robust Comparison Engine

- Data Comparison – even using different data extraction rules for each document
- Image Comparison – can compare a pie chart in Metacode to that from AFP – even though they are drawn very different technically
- Text Comparison – can ensure that all the same words are used, even if paragraph flow is different – perhaps due to font-size change, or text box size
StreamSure Features...

- Batch and GUI-Mode Compare
- Decoded Print-Stream Comparison – for the print-shop floor, useful to see every little change that might have occurred.
  - Much faster than windiff – as StreamSure compares one page at a time! RAM and CPU time is significantly reduced compared to those tools
Demonstration Time

Seeing StreamSure in Action

Task: Compare two documents that are “the same” – but are they “the same”? Why? Why not?