

# Publishing Automation: What Can You Do Today?

---

Thad McIlroy  
Arcadia House  
San Francisco & Toronto

Presented to  
Gilbane Boston  
November 28, 2006

# My Background 11/2006

- 8 years in bookselling & publishing in Canada; 4 in the U.S.
- 18 years studying the intersection of technology and print publishing, working with both printers & vendors
- 5 years with Seybold Seminars

# More Recent Background

- 10 years studying the impact of the Internet on graphic communications
- Major focus now:
  - The future of publishing (.com) (year-end)
  - Workflow (JDF)
  - Publishing automation

**We Need  
a Little Pain  
Assessment!**

## PAIN ASSESSMENT CHART

0	2	4	6	8	10
Sin Dolor	Leve	Moderado	Severo	Muy Severo	Intolerable
不痛	輕微的痛	有點痛	很痛	非常痛	無法忍耐的痛
Walang Sakit	Masakit Kunti	Medyo Masakit	Masakit Talaga	Napakasakit	Sobrang Sakit
He bonut	Cherica bonut	Spaghad bonut	Umbho bonut	Oranoculano bonut	Helenoculano bonut
No Pain	Mild	Moderate	Severe	Very Severe	Exeruciating
					

# What is Automated Publishing?

- 20 minutes to publish...
- After 4 months of preparation!

# What is Automated Publishing?

- Not “page by page” layout (duh!)
- Template-based, and/or
- Rules-based
- Able to react to conditional circumstances (i.e., illustrations, marginalia, etc.)
- “Tweakable”
- Cross-media

# Cross-Media Issues

- What is cross-media?
- Is all media of equal importance?
- Where does the content originate?
- Should ALL content appear in each medium?



# What Happened to *Batch*?

- Intended for relatively simple page designs with a large number of repeating elements (directories, dictionaries, etc.)
- Limited ability to deal with graphics, especially color
- Varied typographic limitations

# A Product History of AP

- Typesetting systems
- Techdoc systems (Interleaf, Frame)
- TeX and LaTeX
- SGML-based systems
- Desktop page layout with style sheets and plug-ins
- The next generation (largely XML-based)

# Personalization

# The Web

# Technology Requirements

*An Understanding of...*

- XML (& XSL, XSL, XSLT, etc. etc.)
- DTDs and Schema
- DAM & CMS
- Accessibility
- Page production
- Typography
- Etc., etc.

# Complexity

**Gilbane Report Email News Vol. 8, Num. 39 - Message (HTML)**

File Edit View Insert Format Tools Actions Help Adobe PDF

Reply Reply to All Forward

From: NewsShark [newsshark@gilbane.com] Sent: Mon 11/20/2006 8:20 AM  
To: thad@arcadishouse.com  
Cc:  
Subject: Gilbane Report Email News Vol. 8, Num. 39

an alphabetical list, or make sure that top customer support questions, product details or sales tools are updated in real time based on the "wisdom of the crowd," or measures of popularity, such as most requested pages. As an on-demand, web-based solution, WebSideStory Publish 4.0 puts control of web site content in the hands of the content owners themselves. With the launch, WebSideStory completes the integration between the different components of its on-demand digital marketing suite, including web analytics (HBX), site search, web content management and keyword bid management. WebSideStory Publish 4.0 enables marketers to integrate content from other sources into their site, including RSS feeds, database information, news articles and data from mapping services. WebSideStory Publish 4.0 allows for more sophisticated content creation. In addition to enabling analytics-driven content, this new system can automatically build a site map, generate site navigation based on the architecture of a site, or make sure content meets certain criteria, such as requiring FAQs to end in question marks. <http://www.websidestory.com>

**Idiom Enhances WorldServer Product Family**  
11/14/2006

Idiom Technologies, Inc. announced enhancements to its WorldServer product family, including new versions of the company's WorldServer globalization management system (GMS) and WorldServer Desktop Workbench, a Windows-based desktop computer aided translation (CAT) tool. The new version of the WorldServer GMS platform offers support for Japanese as a source language. The new version of WorldServer Desktop Workbench features open standards support, tighter integration with WorldServer and better overall usability. WorldServer 8.0.1 new highlights include: Support for Japanese source language Made possible through the introduction of a new word-breaking algorithm that handles Japanese source text; Fully Japanese localized user interface and online help; WorldServer users have a new Personal Preference that allows them to control the "Display Language" of the WorldServer user interface. Each user in the system can have his/her own personal preference, allowing some users to work with the system in English while others work in Japanese. WorldServer Desktop Workbench 8.0 new highlights include: Support for Microsoft Office Spell Checking; Support for XLIFF- and TBX-based translation kits - WorldServer will now offer support for a new translation kit format called XLZ. Desktop Workbench 8.0 includes new features that provide tighter integration between the Workbench and WorldServer; Idiom has worked with several translation companies to improve the usability and efficiency of Desktop Workbench. <http://www.idiominc.com>

**Endeca Partners with Clarabridge**  
11/14/2006

Endeca announced a partnership with Clarabridge to expand and simplify integration between the Endeca Information Access Platform (IAP) and business intelligence (BI) tools. According to the terms of the deal, Endeca will resell Clarabridge BI Connectors for Business Objects, Cognos and MicroStrategy BI platforms as value-added components of the Endeca IAP. The combined offering will help Endeca IAP customers extract additional value from BI investments, unite reports and metrics with other application data and content sources, and extend business intelligence to a wide audience of users. The Clarabridge Connectors give organizations new abilities to index and make available data and reports held within the most widely used business intelligence platforms. <http://www.clarabridge.com>, <http://www.endeca.com>

**PTC Announces Arbortext 5.3 with Additional Support for DITA**  
11/13/2006

PTC (Nasdaq: PMTC) announced the newest version of its dynamic publishing software, Arbortext 5.3. This version extends PTC's support for the Darwin Information Typing Architecture (DITA). DITA, an OASIS-sponsored initiative, is an XML-based architecture for authoring, producing, and delivering technical information. IBM donated DITA to OASIS in March of 2004, where it is now managed by the OASIS DITA Technical Committee. PTC believes that by the end of 2008, up to 80% of all new XML publishing installations will be based on DITA. DITA

# Paging Rules

1. Facing pages are equal in length. Long and short spreads are allowed, but only on facing pages.
2. Widows over  $\frac{3}{4}$  line are permitted below illustrations, and below tables.
3. Spaces above all separate line heads (but not below) may be increased in paging. Spaces around extracts, mathematics, lists, illustrations, etc. may also be increased in paging but should not be decreased.
4. Space above separate line heads and space-break indicators (such as asterisks) at the top of a new page is not retained.
5. Space below chapter titles may be varied if necessary.

# Paging Rules, Pt. 2

6. No less than two lines of text are permitted below subheads at foot of page. A minimum of four lines of text is required on any page with illustration(s) or table(s), at the end of a chapter.
7. Maximum number of consecutive line-end hyphens is three.
8. Turned-page tables and illustrations read correctly when the book is turned to the right.
9. Pages that are at the end of chapters or sections should have at least 6 lines of text on them.



# Design

- There will be a constant tension between design sophistication/complexity and production productivity
- They occupy very different positions in the value chain — different cultures; different incentives
- There's an inbred bias against automation in much of the design community

# Templates & Automated Publishing

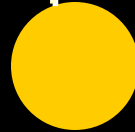
- Page designs must be translated into programmable templates
- While design is creative, template creation is essentially “engineering”
- Design is supposedly rules-based, but breaking the rules is highly valued
- Templates are about repeatable rules; the more exceptions, the less the efficiency

# Islands of Automation

authoring &  
editing



illustration &  
photography



rights



production &  
preflight



distribution



# Overview of Workflow

1. Project initiation/planning
2. Content development
3. Editorial
4. Production/composition
5. Printing
6. Publishing

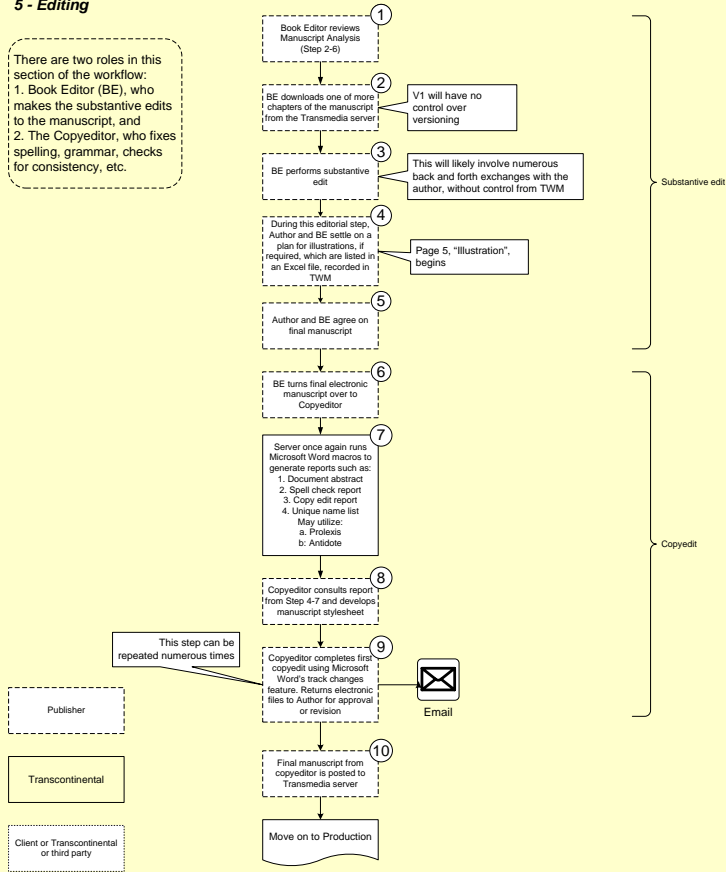
# Functional Areas

- Composition/page make-up
- Project management
- Text printer prep
- Art rendering
- Copyediting
- Proofreading
- Page proofs
- Image permissions & licensing
- Air freight/couriers
- Cover design/production
- Web-related
- Interior/interface design
- Digital/film proofs

# Lots of Flow Charts

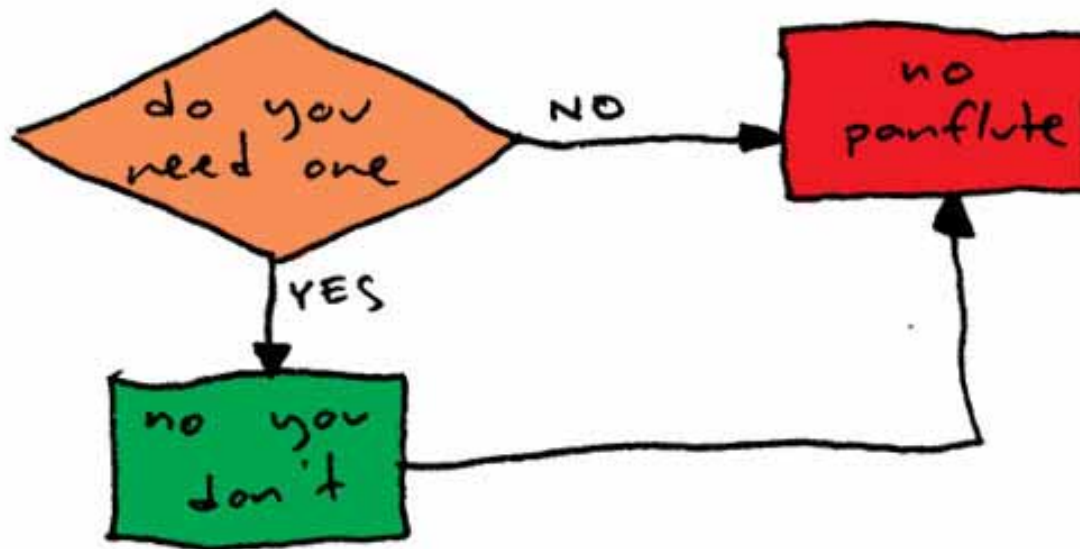
## 5 - Editing

There are two roles in this section of the workflow:  
 1. Book Editor (BE), who makes the substantive edits to the manuscript, and  
 2. The Copyeditor, who fixes spelling, grammar, checks for consistency, etc.



# My Favorite Flow Chart

PANFLUTE FLOWCHART



# Tools

## High-end

- XyEnterprise XPP
- Datalogics DL Pager and DL Composer
- PTC/Arbortext

## Tech Doc

- Adobe's FrameMaker

## Specialized

- Antenna House & RenderX XSL Formatters
- TeX and LaTeX



# Tools (2)

## Quark/InDesign-based

- KyTek Autopage
- Typéfi
- MEI
- WoodWing

## POD/VDP

- XMPie
- PageFlex

# This Conference

- **AP-2: A Tale of Three Successful Publishing Automation Projects**
- **AP-3: How to Develop a Solid Publishing Automation Plan**
- **AP-4: Integrating Automated Publishing into the Broader DAM and CMS Process**
- **AP-5: Five Minutes with a Publishing Automation Vendor**

**Q&A**

Thank you

[thad@arcadiahouse.com](mailto:thad@arcadiahouse.com)