Managing Knowledge in the Era of the Web

Thad McIlroy
President, Arcadia House
October, 2002

Outline

- **The nature of the impact of the Web**
- **The author's changing role**
- **Asset management and content management**
- **Workflow**
- **Format and structure**
- **PDF**
- **The human factor**

My Background

- **Working in publishing since 1977**
- **Consulting to publishers about technology since 1985**
- **Five years at Seybold Seminars**
- **A gazillion articles and seminars on digital technology for publishing**
- **www.arcadiahouse.com**

How Has the Internet Changed the Role of the Author?

Role Change

If you had to choose just one:

1. It **has** changed their role completely
2. It **hasn't** changed their role at all

The Struggle

The struggle, as always, is between content and form,
between commerce and creativity,
between access to capital & distribution and smaller, independent voices.

The Internet Marches On

- **"49% (5.8 million of 12 million households) had at least one member that regularly used the Internet at home in 2001."**

- **Up 1.1 million from 2000.**

  — The Globe & Mail, July 26, 2002

The Internet and Publishing

- **The potential impact is on every part of the publishing process:**
    - Content and form
    - Commerce and creativity
    - Corporate concentration vs. independent voices
- **How it actually plays out is another matter!**

Love Those New Tools!

"I mistrust this computer work. It's too clean. The result is coming too quickly. I write all manuscripts still and first of all by hand, with pen and paper..."

— Günter Grass

Winner, Nobel Prize for Literature, 1999

First, a Simple Distinction

1. **Fiction**
2. **Non-Fiction**

Let's focus just on non-fiction, and call it (generalizing broadly):

"Information Publishing"

How Has Digitization and the Web Changed Information Publishing?

Completely!

And in every way…

■ In more detail…

The Process of Publishing

- **Authoring**
- **Production**
- **Distribution**

The Process of Publishing

- **Authoring**
  - **research**
  - **writing**
  - **assembly of illustrations**
- **Production**
  - **editing**
  - **composition**
- **Distribution**
  - **marketing**
  - **physical distribution**

The Process of Publishing

- **Authoring**
  - **research**
  - **writing**
  - **assembly of illustrations**
- **Production**
  - **editing**
  - **composition**
- **Distribution**
  - **marketing**
  - **physical distribution**

Digitization and the Web
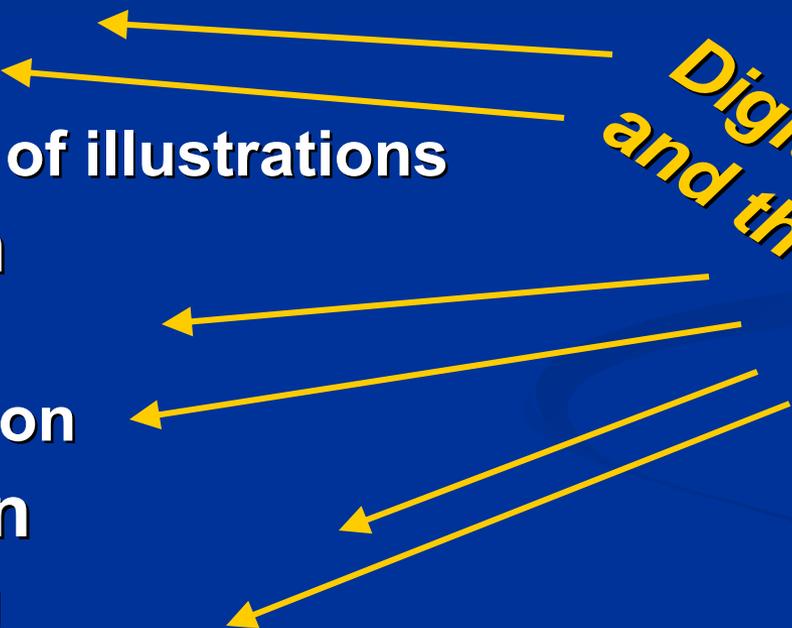
This is Our Biggest Yet

"When desktop publishing was launched into the world in 1984 it contained the shock of the new, and we thought we were witnessing a revolution. It was only later, when things settled down a little, that we could see that all we had achieved was the digitization of analog processes. We were still doing the same things for the same reasons, but with different tools; nothing had really changed."

— The Columbia Guide to Digital Publishing

The Move to Content Management

- **The move from being exclusively print-focused…**

- **To encompassing multiple media — print, the Web, CD-ROM, wireless…**

- **And managing that production digitally**

Asset Management
and
Content Management

Terminology

Digital Assets:
Electronic files containing text, images, sounds and/or moving images.[1]

1. As well as data about the assets (metadata).

Terminology

Asset Management:
Software system used to catalog digital content for archiving or reuse.

Terminology

Content Management:
Technology and processes to maintain digital assets in a form where they can be quickly and easily reclaimed and reused in publishing workflows.

What is Content Management?

Collecting, managing, and making content[1] available

1. Content = all digital assets: text, images, rich media, and metadata

Content Management vs. Asset Management

- **Asset management just stores content**
- **Content management *deploys* content into multiple media**
- **Content management = asset management + workflow**

Terminology

Workflow:

The sum of the processes involved in the creation of print and/or electronic pages.

The Workflow Problem

- **Product creation and production control functions are shared by the author, the publisher *and* printer — at different locations, at different times, and with different priorities**

The Value of Content Management

- **Control assets**
- **Improve workflow**
- **Control rights and permissions**
- **Publish to multiple media**
- **Reduce time to market**
- **Offer new products from the same content**

On the other hand, it can be expensive...

Figure 5 Basic Costs Of Content Management

Investment	Type	Cost	Vendors
Content management software	License	$250,000	Vignette, Documentum, Interwoven
Professional services	Implementation* Training	$180,000 $3,000[†]	Software professional services groups plus integrators like Sapient, IBM Global Services
Knowledge engineering	Controlled vocabulary and content hierarchy	$200,000[‡]	Requisite, SAQQARA, Metacode (Interwoven)
Total:		**$633,000**	

*Three developers for 8 weeks @ $1,500/day
[†]Two-day training course for 3 staff members @ $1,000/pupil
[‡]Average cost for 50,000 pages

Source: Forrester Research

Becoming a Content-centric Business

- **The existing model is "product-centric"**
- **Clearly the *form* of the products is fast-changing**
- **But the *content* has enduring value**
- **The publisher's core competency is content-creation and distribution**
- **And it's existing customer relationships**

The Importance of XML

- **XML fundamentally enables content management**
- **Combining of the power of style sheets with the power of databases**
- **Style sheets with meaning**

Format vs. Structure

- *Format* describes how content is intended to look when it is displayed or printed

- *Structure* describes the purpose or meaning of content

A Case Study

- **The Columbia Guide to Digital Publishing**
- **Author templates with full tagging for structure and format**
- **Editing handled in Microsoft Word**
- **Direct tagged output**
- **Online and print version**

PDF

- **The perfect "tweener" format**
- **Appealing to publishers in between the all-paper generation and the upcoming all- (or mostly all-) electronic generation.**

PDF for Publishing

- **Authoring**
  - research
  - writing   PDF
  - assembly of non-text materials
- **Production**
  - editing   PDF
  - composition   PDF
- **Distribution**
  - marketing   PDF
  - physical distribution   PDF

The Human Factor

- **This is complex and sophisticated technology**
  - **Far tougher than DTP/prepress**
- **Implementation takes time**
- **Plan to study and learn gradually**
- **The potential benefits of CMS are enormous!**

Thank you

thad@arcadiahouse.com