

# The Dynamics of Internet Publishing on the Computer Book Industry

By:  
Robbie Allen

Submitted to the System Design and Management Program and the Department  
of Civil and Environmental Engineering in Partial Fulfillment of the Requirements  
for the Degrees of

**Master of Science in Engineering and Management**

and

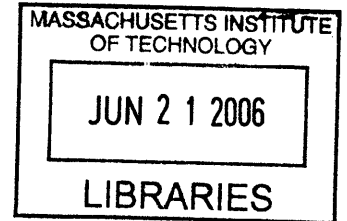
**Master of Science in Civil and Environmental Engineering**

at the

Massachusetts Institute of Technology

June 2006

© 2006 Robbie Allen, All Rights Reserved



The author hereby grants to MIT permission to reproduce and distribute publicly paper and  
electronic copies of this thesis document in whole or in part.

**BARKER**

Author: .....  
Robbie Allen  
System Design and Management Program

Certified by: .....  
Thesis Supervisor: Prof. Michael Cusumano  
Sloan Management Review Distinguished Professor, MIT Sloan School of  
Management

Certified by: .....  
Thesis Supervisor: Prof. John Williams  
Director, Auto-ID Lab, and Professor, Civil and Environmental Engineering

Certified by: .....  
Pat Hale  
Director, System Design and Management

# Table of Contents

<b>List of Figures</b> .....	<b>4</b>
<b>List of Tables</b> .....	<b>5</b>
<b>Abstract</b> .....	<b>6</b>
<b>Acknowledgements</b> .....	<b>7</b>
<b>1. Introduction</b> .....	<b>8</b>
1.1 Internet Publishing vs. Traditional Publishing.....	9
1.2 Thesis Purpose and Structure.....	10
<b>2. Analysis of the Computer Book Industry</b> .....	<b>11</b>
2.1 State of the Computer Book Market.....	12
2.2 Limitations of Traditional Publishing.....	15
2.3 Summary .....	22
<b>3. Internet Publishing Opportunities and Challenges</b> .....	<b>24</b>
3.1 Advantages of Internet Publishing.....	24
3.2 Challenges with Internet Publishing .....	30
3.3 Internet Publishing Models .....	35
3.4 Summary .....	42
<b>4. Exploration of Business Models</b> .....	<b>43</b>
4.1 Free .....	43
4.2 Advertising.....	44
4.3 Subscription .....	44
4.4 Unit .....	45
4.5 Sponsored.....	46
4.6 Earn.....	47
4.7 Summary .....	47
<b>5. Improving the Content Development Process</b> .....	<b>48</b>
5.1 Comparison to Software Development.....	48
5.2 Applying Software Development Concepts.....	51

5.3 Summary .....	57
<b>6. Conclusions.....</b>	<b>58</b>
6.1 Important findings.....	58
6.2 Future research.....	59
<b>Bibliography .....</b>	<b>61</b>

## List of Figures

Figure 1. Weekly Bookscan Trends from 2003-2006.....	13
Figure 2. Waterfall Development Lifecycle .....	16
Figure 3. Anatomy of the Long Tail .....	28
Figure 4. U.S. Daily Newspaper Circulation versus Number of Households.....	36
Figure 5. Cumulative Weblogs, 2003-2006.....	39
Figure 6. Estimate-Convergence Graph.....	54

## List of Tables

Table 1. Top nine computer book publishers as of April 09, 2006 .....	12
Table 2. Top forty computer books as of April 09, 2006 .....	15

## **Abstract**

The Internet has been a disruptive force for many industries, but perhaps none more so than in the publishing business. While many segments of the publishing industry have made attempts to use the Internet to augment or replace existing revenue sources, none has done as little with the Internet as book publishers. This thesis will examine the computer book niche of the publishing business, review various internet publishing models that have been employed to date, and outline opportunities and challenges that computer book publishers should consider to stay viable in the Internet age. An analysis of various internet publishing business models will be covered along with a discussion of how internet publishing can facilitate better methods and processes for developing content.

Thesis Supervisor: Michael Cusumano

Title: Professor, MIT Sloan School of Management

Thesis Reader: John Williams

Title: Professor, Civil and Environmental Engineering

## **Acknowledgements**

Like most achievements in my life, completion of this thesis would not have been possible without the support of my family and colleagues. My wife Janet was supportive throughout my MIT journey despite the sacrifices she had to make while I was away. For that I will always be grateful. Thanks to my thesis advisors, Michael Cusumano and John Williams, for their guidance and patience. Professor Cusumano's course on the Software Business was one of the highlights of my time at MIT. My sincere appreciation goes to Pat Hale, Director of SDM, for his support and flexibility. The SDM program is successful in large part due to Pat's leadership and empathy for busy professionals.

# 1. Introduction

Over the past ten years, the publishing industry has been forced to deal with a new technology that has had an impact the likes of which hasn't been seen since the invention of the Gutenberg press in 1450. That technology is the Internet, and it has changed everything from the way consumers think about finding and reading content to how authors and publishers create and sell content.

The publishing industry has roots dating back to 2400 B.C. when papyrus sheets were glued together to form scrolls [1]. But it wasn't until the invention of the printing press in the fifteenth century that publishers could distribute printed materials in large quantities. Fifty years after Gutenberg's printing press, the number of books published per year went from a few thousand to over 9 million [2]. Today, the book publishing industry in the United States alone consists of over 2,600 companies with combined annual revenue of \$30 billion [3]. Since 1776, over 22 million books have been published in the US with over a billion copies in circulation [4].

The publishing industry is known for being slow to adopt change, especially with technology-related advances. With the wide-scale proliferation of the Internet, the publishing landscape has been altered dramatically. Now, information can be packaged and distributed in ways never before imagined. Printed materials are no longer the primary or even the preferred way to communicate information. Restricted by their current business models and disinclined to understand the significance due to the inertia of change, book publishers have made only incremental improvements leveraging the Internet instead of the significant advances that are possible.

The Internet represents a sea change to the way content is developed, produced, distributed, and sold. This thesis will examine the opportunities and challenges as content moves online and explore ways in which computer book publishers can adapt to remain viable in the Internet age.



## ***1.1 Internet Publishing vs. Traditional Publishing***

A distinction is made in this thesis between traditional publishing (sometimes referred to as print publishing) and internet publishing. In this context, traditional publishing refers to publishers of printed materials such as books, magazines, and newspapers. Internet publishing refers to publishers whose focus is delivering content via the Internet.

Most traditional publishers today make use of the Internet in one form or another, but do not rely on it as a significant source of revenue. Traditional publishers have made only half hearted attempts to embrace the Internet because it presents such a different way of doing business. Also, due to the open and free flowing nature of the Internet that the public has come to expect, most publishers are unsure how to maintain the same profitability levels compared to the print business.

The Internet is changing all forms of traditional publishing from books and magazines to newspapers and journals as people look to the web for information. This thesis will target a subset of the traditional publishing industry, specifically the computer book business. Computer book publishing makes for a good case study because computer book publishers have to be some of the most agile and adaptive in the publishing business given the nature of the topics they develop content for. Computer book publishers are finding the Internet to be their top competitor as content that would typically be found in printed computer manuals is now available online (for free) in many cases.

Many of the issues covered in this thesis apply to other types of publishing including non-fiction and fiction books as well as printed magazines and newspapers. However, these sectors have their own set of issues with the way content is developed, marketed, sold, and consumed that make them different enough so that broad generalizations cannot be made across the publishing industry without significantly more research. On average, the customer base for computer books tends to be very technically savvy and uses the Internet on a daily basis. That combined with the fact that computer book publishers often develop content about the Internet make them a great candidate to use the Internet as a content delivery platform.

## ***1.2 Thesis Purpose and Structure***

The research for this thesis was conducted over an eight month period. A combination of primary literature, conversations and informal interviews with computer publishers and editors, and the author's personal experience as a computer book editor and author were used to inform the structure, analysis, and recommendations presented here.

The primary purpose of this thesis is to analyze the computer book industry and explore viable alternatives to print publishing using the Internet as a major content delivery platform. Business models and content development processes will be explored and recommendations will be made for opportunities the Internet offers as well as challenges pertaining to moving content online.

In Chapter 2, the computer book publishing industry will be analyzed including the growth of the business (or lack thereof) and major challenges the industry faces.

Chapter 3 covers the advantages of internet publishing from a process perspective and describes how the limitations described in Chapter 2 can be addressed. Next, the challenges with internet publishing are outlined to show some of the difficulties in making it a feasible model. Lastly, the different forms of internet publishing that have been used ranging from content websites to blog networks will be reviewed.

The biggest question when it comes to creating an internet publishing business is how to make money at it. Chapter 4 explores this question and reviews what has and hasn't worked so far and suggests approaches computer book publishers can take.

Chapter 5 looks at the methods for developing content and how the Internet presents new opportunities to improve decades old processes. A comparison is made between software engineering and content engineering in order to draw parallels so that lessons learned from the former can be applied to the latter.

Chapter 6 summarizes the major finding from this thesis and highlights several recommendations that those working in computer book publishing should consider in order to stay competitive.

## **2. Analysis of the Computer Book Industry**

As computers have become more popular, the need for computer-related books has grown in a similar pattern. Large publishing firms have gone from releasing individual computer titles to creating specific imprints dedicated to computer books. By the early 1980s, the first independent firms devoted to computer books were established. In the late 1990s, there was a significant increase in the number of computer book publishers as more firms tried to take advantage of the Internet boom. Since the downturn in 2001, there has been considerable consolidation as sales dropped by over 50% for many publishers.

Detailed sales data for books in the US has been tracked by Nielsen Bookscan since 2001 [5]. Publishers and distributors can subscribe to the Bookscan service to receive weekly reports on sales data from over 4,500 retailers in the US. Bookscan data is broken out by segment. Computer book publishers subscribe to the "Top 10,000 Computer Books" service to receive information on only computer-related titles.

As of the April 2006 Nielsen Bookscan report, there were 156 firms listed that sold 10 or more units per week in the computer book sector. Many of these firms are nothing more than publishing imprints that belong to larger publishing houses. The publishing industry is notoriously incestuous with large conglomerates owning several, sometimes competing, labels.

Table 1 shows the top nine computer book publishers from the April 09, 2006 Bookscan report.

<b>Publisher</b>	<b>Wk Units</b>	<b>Market Share</b>	<b>YTD Units</b>	<b>Market Share</b>
For Dummies	22,533	15.27%	384,519	17.09%
O'Reilly	20,858	14.14%	303,247	13.48%
Microsoft	14,655	9.93%	212,958	9.47%
Peachpit Press	11,863	8.04%	205,714	9.14%
Que	8,247	5.59%	122,962	5.47%
John Wiley	7,449	5.05%	115,980	5.16%
McGraw-Hill/Osborne	7,218	4.89%	112,018	4.98%
Sams	6,073	4.12%	88,260	3.92%
Addison Wesley	4,984	3.38%	69,357	3.08%

**Table 1. Top nine computer book publishers as of April 09, 2006**

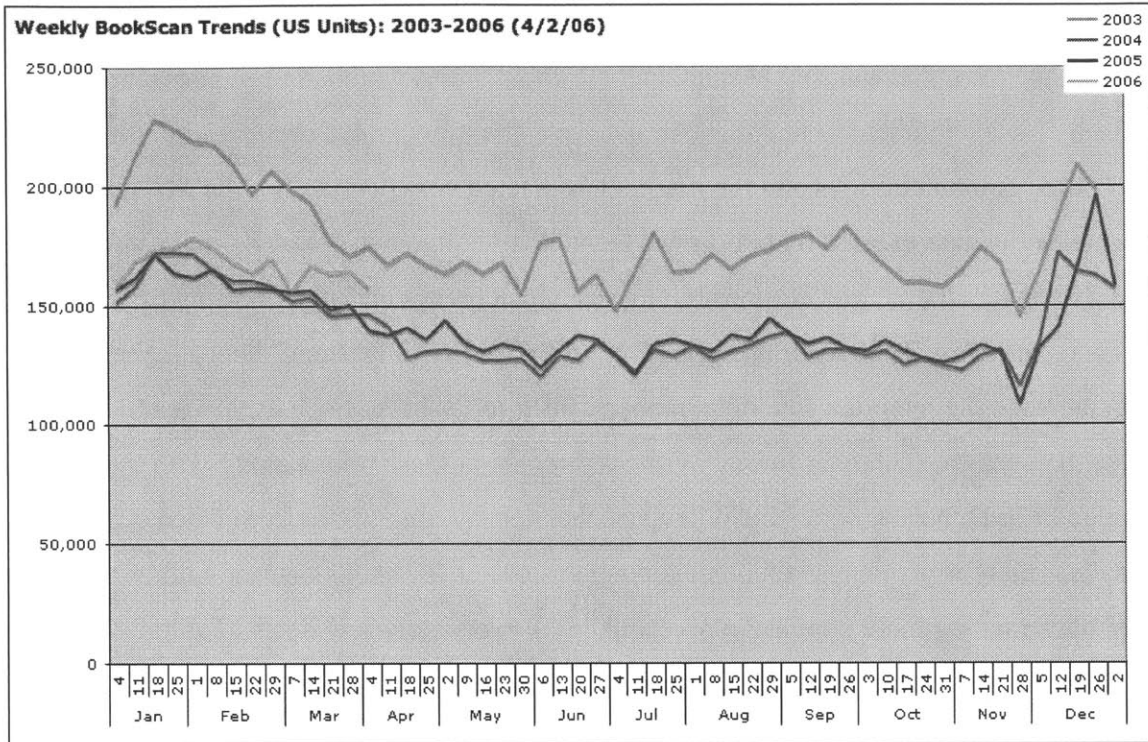
For Dummies and John Wiley are part of Wiley Publishing. Addison Wesley, Sams, Que, and Peachpit Press are part of Pearson Education. McGraw-Hill, Microsoft Press, and O'Reilly are independent firms. That means five companies make up the top nine publishers and have close to three quarters of the overall market share.

## ***2.1 State of the Computer Book Market***

The market for computer books closely follows the demand for information technology. There is a direct correlation between the technologies that are rising in popularity and sales for books that cover those technologies [6]. The peak of the computer book industry can be mapped with the peak of the Internet boom. Computer book sales experienced 20% year-over-year declines from 2002-2004 [7]. 2005 showed a slight increase, but is still far off from the 2001 highs. As IT departments put more money back into deploying new technologies, IT professionals put more money into computer books to become educated on those new technologies.

Tim O'Reilly, the CEO of O'Reilly Media, has been at the forefront of using Bookscan data to analyze the computer book market. He's created a research team at O'Reilly Media that loads the weekly Bookscan data into a MySQL data mart in order to do reporting and historical trending.

The following graph was created by O'Reilly Research in April 2006 that shows yearly comparison of book sales back to 2003 [7]:



**Figure 1. Weekly Bookscan Trends from 2003-2006**

The figure shows that the computer book business is slowly crawling out of the Internet winter that hit the entire IT industry, but sales levels still aren't close to what they were four years ago. The computer book market is still down over 40% from its highs in 2001.

### 2.1.1 Industry Outlook

Despite the recent growth in computer book sales, it is unlikely that the industry will ever see the same level of profitability that it did before 2002. One underlying trend for the decreased sales is increased quality and depth of technical documentation on the Internet. Many IT professionals turn to the Internet first to learn about a new topic or to answer questions before reaching for a book. For example, the Microsoft Developer's Network (MSDN) has slowly improved over the years to the point where it offers high quality and

in-depth documentation that is comparable to most books. As more people use these sites to learn about technologies, the market for reference-oriented computer books diminishes.

Despite the sluggish growth of the overall industry a few books have been able to flourish during this time. The Missing Manual and Headfirst series which are both developed by O'Reilly have continued to churn out bestsellers. These books are more entertaining than traditional computer books and the market has reacted very positively. The Missing Manual series was created by NY Times technology columnist David Pogue. His writing style is conversational and at times humorous which breaks up the monotony of learning a new technology. The Headfirst series employs a variety of learning theory techniques to help increase the retention rate of its readers. Both series have much higher entertainment value than typical computer books. With the success of these books along with the much heralded For Dummies series by Wiley and the Step-by-Step series from Microsoft Press, it is clear there is growing preference among customers for books that are highly readable over ones that are more comprehensive and cover every technical detail. Table 2 highlights this point. Twenty of the top forty best selling books as of the April 2006 Bookscan report came from four series (For Dummies, Step-by-Step, Missing Manual, and Headfirst).

<b>Rank</b>	<b>Publisher</b>	<b>Title</b>
1	For Dummies	Windows XP For Dummies
2	Peachpit Press	iPod Book: Doing Cool Stuff with the iPod and the iTunes Music Store
3	For Dummies	Excel 2003 for Dummies
4	Peachpit Press	The iPod & iTunes Pocket Guide
5	O'Reilly	Mac OS X: The Missing Manual, Tiger Ed
6	For Dummies	eBay for Dummies
7	Microsoft	Microsoft Office Excel 2003 Step by Step
8	For Dummies	Windows XP All-in-One Desk Reference For Dummies
9	For Dummies	PCs For Dummies
10	Microsoft	Microsoft Office Project 2003 Step by Step
11	McGraw-Hill/Osborne	QuickBooks 2006: The Official Guide
12	Macromedia Press	Macromedia Dreamweaver 8: Training from the Source
13	Adobe Press	Adobe Photoshop CS2 Classroom in a Book
14	For Dummies	Excel 2003 All-in-One Desk Reference for Dummies
15	New Riders Pragmatic	The Photoshop CS2 Book for Digital Photographers
16	Programmers	Agile Web Development with Rails: A Pragmatic Guide
17	Peachpit Press	HTML for the World Wide Web with XHTML and CSS:

		Visual QuickStart Guide
18	Que	Upgrading and Repairing PCs
19	Microsoft	Microsoft Office Access 2003 Step by Step
20	New Riders	The Photoshop Elements 4 Book for Digital Photographers
21	For Dummies	QuickBooks 2006 For Dummies
22	Peachpit Press	The Photoshop Channels Book
23	For Dummies	The Internet For Dummies
24	For Dummies	Office 2003 All-in-One Desk Reference for Dummies
25	Manning	Ajax in Action
26	For Dummies	Access 2003 All-in-One Desk Reference for Dummies
27	O'Reilly	Head First Design Patterns
28	Gotham Books	The Official eBay Bible
29	RMC	PMP Exam Prep: Rita's Course in a Book for Passing the PMP Exam
30	Rough Guides	The Rough Guide to iPods, iTunes & Music Online -
31	John Wiley	Excel 2003 Bible
32	O'Reilly	Head First HTML with CSS & XHTML
33	For Dummies	Digital Photography For Dummies
34	O'Reilly	Adobe Photoshop CS2 One-on-One
35	Microsoft	Microsoft SQL Server
36	For Dummies Pragmatic	Word 2003 for Dummies
37	Programmers	Programming Ruby: The Pragmatic Programmers' Guide
38	Microsoft	Microsoft Visual Basic 2005 Step by Step
39	For Dummies	PowerPoint 2003 for Dummies
40	Peachpit Press	Macromedia Dreamweaver 8 Hands-On Training

**Table 2. Top forty computer books as of April 09, 2006**

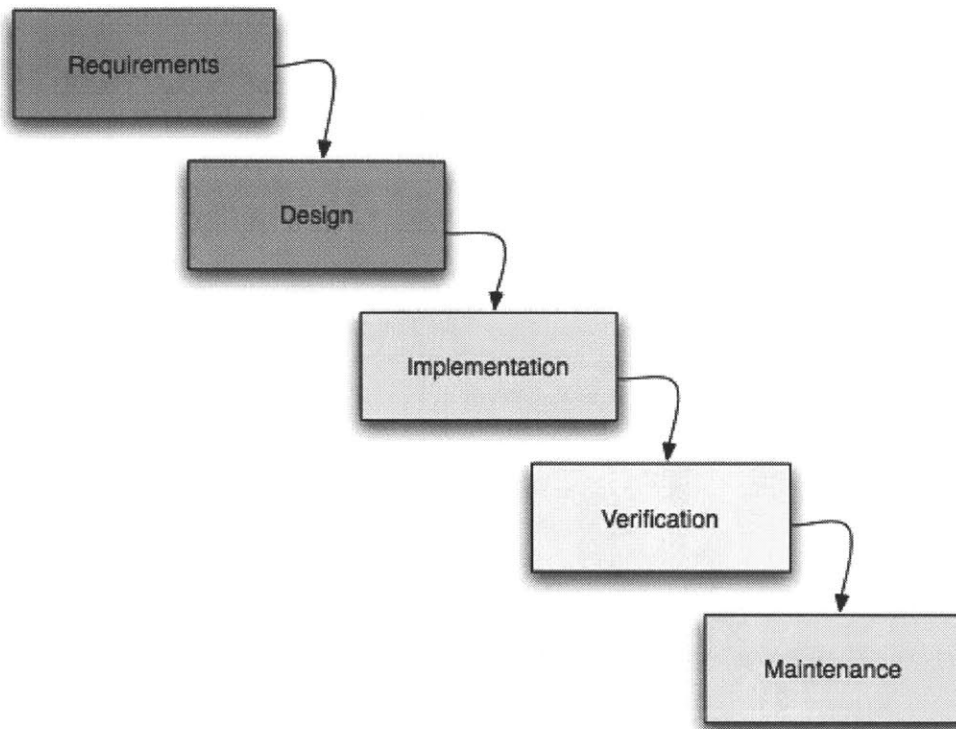
## ***2.2 Limitations of Traditional Publishing***

It is important to look at the current limitations of the computer book publishing industry and question all assumptions about how content is created and consumed. The Internet offers a blank canvas on which to create new forms of content and new ways to conduct business. We must look deep within the publishing industry in order to discover possibilities that those entrenched in the field take for granted.

### **2.2.1 Restrictions of a printed medium**

Perhaps the biggest problem with print publishing is that by definition it operates in the physical world. Producing books has many of the same limitations as producing any physical good. This is one of the big dichotomies between printed content and Internet content. Developing printed content is analogous to developing computer hardware

whereas developing Internet content is analogous to developing computer software. Physical goods tend to follow a waterfall-style development process [9], which has many limitations. Figure 2 shows the steps involved in the waterfall process [9].



**Figure 2. Waterfall Development Lifecycle**

As you can see from the figure, there are no iterations in the process. In practice there may be some limited iterations and feedback loops but typically there are none based on actual customer feedback. Generally, real-world customers are not involved in the computer book writing process at all. This is different from other forms of book publishing, such as business books or fiction, where authors may serialize their work in magazine or journal articles before putting the work in a book. As a rule, computer book authors are always crunched for time as the clock ticks on the underlying technology they write about.

Another artifact of printing books is that there has to be something substantial before it can be released to the public. Just as you can't ship half of a computer processor, you



can't ship half of a book; it is all or nothing. This elongates the development process. The content for chapter 1 of a book may be ready on day 30, but it is not until all other chapters are completed before the book will be printed (which can take 1 to 2 years). There is a huge opportunity cost as the early material lies dormant.

Another issue that has to be carefully considered when dealing with physical goods is inventory and distribution costs. Publishers have to make a significant investment in each book and a certain number of copies have to be sold to recoup these costs. This limits the frequency in which books can be updated so it is usually 2-4 years before new editions are developed.

### **2.2.2 Book is the unit of content**

As discussed in the previous section, publishers cannot print content incrementally. This gives rise to the concept of a "book", which is simply an organizational device for grouping content together that is substantial enough to print. Publishers have always thought in terms of books. Their whole business is built around them. Even the more technically advanced computer book publishers today are still focused on books--whether it is printing them or making them available in electronic format.

The problem is that the purpose of a book loses its usefulness in the context of the Internet. While it is still important to group content together online in meaningful ways, it is not necessary to think in terms of a "book". The Internet offers a great opportunity to redefine our notions of how content is organized. The unit of content should be the smallest piece of information that is useful. By focusing on smaller units of content, it will be possible to remix content in ways never before imagined. This is especially true of task-oriented and referenced-based technical content where much of the material is not dependent on other "chapters".

Tutorials and text that is heavily narrative are more challenging to think outside the boundaries of a book. However, very few people read computer books cover to cover.

Even in the case of tutorials, there are opportunities to divvy up the content online and give readers exactly what they want, when they want it instead of what the publisher and author think readers want.

Ultimately, the reluctance to think outside the boundaries of a book will be one of the biggest, if not the biggest reason most traditional publishers aren't able to successfully navigate the move to the Internet. The publishers that grasp the significance of the Internet are trying to adapt their current business models and practices to work online instead of looking at it with a fresh perspective.

### **2.2.3 Inequality with the Publisher/Author Relationship**

With the exception of a few best selling authors, publishers hold most of the power in the publishing industry. This has resulted in a less than favorable working environment for authors. This is due to a few reasons, but stems from the fact that publishers have to take on a lot of risk when publishing a book. Publishers pay the editing, production, printing, inventory, marketing, and distribution costs, and have to take a big part of the royalties in order to recoup these costs. Because it is so expensive to publish a book and get it into bookstores, it hasn't been practical for authors to self-publish until recently. Before the advent of the Internet, if an author wanted to publish a book he had to go through a publisher to have any chance of reaching a large audience -- that's if he could find a publisher that was willing to sign the book at all. Publishers have enjoyed complete control over the book publishing process at the detriment of authors.

What value do publishers really bring? Do they add much to the process of writing a book or are they just being rewarded for fronting the capital required to print it? The primary services publishers provide to authors include:

- 1) Editing (Developmental, Proof, Copy)
- 2) Production (Layout/Graphics/Covers/Indexing)
- 3) Printing and Inventory

- 4) Marketing and Branding
- 5) Distribution

Authors can outsource the editing and production work. In fact, many publishers outsource this already, so isn't a source of competitive advantage. The biggest costs for a publisher are for printing and inventory management. These essentially go away with internet publishing. While most technical publishers don't do much marketing for each book, there is something to be said for brand recognition, which has to be earned over time. A brand isn't something an author can outsource. Distribution to stores is also something that isn't cost-effective for authors to buy right now in the traditional publishing model. While it is easy to make content available on the Internet, that doesn't mean it is being widely viewed.

To summarize, the main costs associated with print publishing (editing, production, and inventory) either go away or are minimized in the online world. That leaves marketing and distribution, which are more about establishing a reputation and building relationships than expending capital. Since the capital to produce content online is much cheaper, the stranglehold traditional publishers have had on the industry will start to loosen. Internet publishers will need to focus on creating interesting and unique ways to package and distribute content. This is hard for authors to do individually. Publishers also have to work on improving the author work environment by providing richer information about how an author's content is selling so better informed decisions can be made about creating new content. Instead of being in a comfortable position as a requisite source of capital to publish books, internet publishers will need to focus on providing value-added services to attract authors.

There is still a need for publishers in the online world--not because they serve as a source of capital, but because they provide increased exposure to content and an efficient work environment.

## 2.2.4 Inefficient development process

The way most publishers develop content is surprisingly manual and as a result very inefficient. Computer book publishers with ten or more years under their belt still haven't done much to automate the day-to-day process of developing content. They've automated bits and pieces and some of the production process, but the act of developing content is largely automation free. One publisher that has done a good job at attempting to bring some automation from the software engineering world to publishing is the Pragmatic Bookshelf [10].

The developmental editor of a book is the primary project manager in charge of overseeing the delivery of the manuscript. Editors often use generic multipurpose tools such as Yahoo Groups and Excel to help manage a book's progress. Yahoo Groups is used primarily to facilitate the sharing of files and to keep track of discussions with authors and technical reviewers. Many editors use Excel to track the progress of a book; essentially, a poor man's project management tool. Some people may argue that these tools help keep the process uncomplicated. However, there are many reasons to go beyond simple tools.

One big reason the content development process needs to be automated is to reduce inefficiencies and to streamline the process. There is already a huge opportunity cost with the length of time it takes to produce a book. By making the development process more efficient, publishers could get books out to market quicker which would have a direct impact on sales.

By not automating the process, more responsibility is put on editors. Some editors are good about staying on top of contracts, author submissions, editing duties, and managing tech reviewers, while others are not. Publishers have a hard time distinguishing good editors from the bad ones (from a productivity perspective) until it is too late. If the development process was more automated, publishers could objectively measure the productivity of editors. How many books arrive on schedule? What are the reasons they don't? Is the tech review process taking longer than it should? What is the average time it

takes an editor to edit chapters? All these questions could be answered. Not every book, editor, or author would necessarily need to abide by the same metrics, but by capturing these types of operational statistics, it would be easier to measure productivity objectively.

Automating the process has positive consequences for the business. Getting books done on time is very important because publishers promise booksellers specific dates for each book. When a publisher misses a delivery date for a book, the bookseller often penalizes the publisher by not stocking as many in the stores when it finally becomes available.

Some publishers have systems in place to track the progress of books, but these systems typically have a macro-level view. Automation needs to be driven down to the micro-level, at least as far as monitoring the chapters of a book. An automated workflow system could be put in place so that once an author submits a chapter, a notification is sent to the editor. Once the editor reviews the chapter, a notification is sent to the author to update the chapter or to tech reviewers to review the chapter. All of this could feed a reporting system that provides concise data on the progress of each book. All sorts of interesting reporting could be done to help make informed decisions about adding resources.

### **2.2.5 One-sided conversation**

The Cluetrain Manifesto [23] served as a wake-up call to corporate America about the impact of the Internet on the way people communicate. A central theme of the book is that "markets are conversations" and "through the Internet, people are discovering and inventing new ways to share relevant knowledge with blinding speed". Contrast this with the publishing industry which has been engaged in a one-sided conversation with its customers. Computer book publishers often don't have a clue what customers really want. They throw as many books as they can fund at bookstores in the hope that a few will be successful. Their best attempt to understand customers is reviewing Bookscan numbers and reading Amazon reviews.

One reason the conversation has been so one-sided is because publishers produce only passive learning materials such as books. Some have made half-hearted attempts to provide more interactivity through CDs, but the success of those efforts has been limited.

Whether computer book publishers recognize it or not, they are in the business of educating people. People buy books to either be entertained or to learn. As far as computer books go, especially technical tutorials and references, it is unlikely many people buy them purely for the entertainment value (although an entertaining computer book tends to sell better as discussed earlier). That means publishers should be focusing on optimizing the learning experience for their customers.

Studies have shown [11] that active learning results in significantly better retention than passive learning. Robert Pike, a noted training expert, researched the topic extensively and suggests in his book "Creative Training Techniques Handbook" [24] that we retain:

- 10 percent of what we read;
- 20 percent of what we hear;
- 30 percent of what we see;
- 50 percent of what we hear and see;
- 70 percent of what we say; and
- 90 percent of what we say and do.

That means publishers, who are in the business of educating people, have been focused on techniques that yield a 10% retention rate. By using audio and video technologies via the Internet, retention could be closer to 50%. And going one step further by creating interactive learning experiences, the retention rate could go even higher.

### **2.3 Summary**

The performance of the computer book industry closely follows that of overall IT spending. A few well established publishers have the majority of the market share--a

market that took a nose dive after the dotcom bust. In 2005, the computer book market had the first signs of growth since 2001, but it is still well off its highs. A few series continue to produce best selling books that buck the overall sales trend. Most of these books tend to be more entertaining than the typical computer book.

There are many warts in the computer book publishing industry that were covered in this chapter. The next chapter will evaluate the advantages and disadvantages the Internet offers as an alternative medium to print.

### **3. Internet Publishing Opportunities and Challenges**

This chapter covers the advantages the Internet offers in terms of developing content along with some new opportunities to improve publishing processes that aren't available in the print world. This is balanced with a review of the challenges that internet publishing poses and obstacles that must be overcome in order to be successful at it. The chapter concludes with a discussion of the main internet publishing models that have been used to date along with the pros and cons of applying each to computer book publishing.

#### ***3.1 Advantages of Internet Publishing***

Section 2.2 lists the limitations of the traditional publishing industry, some of which are rooted in decades old publishing practices and others that are based on the innate constraints of print media. Using that as a basis, this section will cover how the Internet can help address those problems and provide new opportunities for publishers to consider.

##### **3.1.1 Content not limited by the pages of a book**

Print materials such as books and magazines are limiting by their very nature. They can't be searched easily. The process of finding content within a book is very inefficient. Publishers have to devote resources to creating an index of key terms and hope everything a reader may want to look up is included. Contrast this with online content, where the latest advances in search technology can be used to allow readers to more quickly find what they are looking for.

Also, by not limiting content to a physical entity, it is possible to remix content in ways not available before. For example, Safari Books Online (<http://safaribooksonline.com>) has created a customized textbook program called SafariU that lets educators mix and match chapters from many different books to create a tailored book for their classes. The next logical progression is to allow the average reader to do the same.



### **3.1.2 Content-as-a-service**

Books are not easy to carry around in even small quantities. In IT, it is common for a programmer or system administrator to use several books at a time when working on a project. Carrying books in between the office and home or to other work sites can be challenging.

Similar to how software has been adapted to the service space (i.e., software-as a-service), the same model could be applied to content. Through the Internet, content can be made available to customers regardless of where they are or when they want it. Many of the same benefits of the software-as-a-service model apply including the ability to provide instant updates and access to content regardless of a user's location.

### **3.1.3 Multimedia and interactive content**

Text is the most common medium for consuming content today, but it is not the only or even the best in terms of retention. Some people prefer to hear the spoken word over reading text. Yet others learn best through visual demonstrations. The Internet allows computer book publishers to move away from text-only publishing and integrate audio and video into the experience. Customers should get to decide what format works best for them.

A recent form of video presentation that has become popular is screencasting. A screencast is simply a video of a computer screen with voice narration that demonstrates the use of an application, tool, or program. Many computer reference books are task-oriented -- they contain a series of instructions on how to accomplish certain computer tasks. This is ideal content for screencasting. Many companies are now starting to use screencasting to provide demos of their products on the web.

### **3.1.4 Apply advances in software development to content development**

There are many similarities between writing code and writing content. Whether the language of choice is Java or English, at the most fundamental level, both are used to express ideas in a structured way. Chapter 5 will cover how the content development process can be improved by applying the things that have been learned in the software development field.

### **3.1.5 Developing content is an ongoing process**

Most computer books are about technologies that are rapidly changing. The books that are first to market often have the best chance for success. As a result, there is a lot of pressure to get books out as quickly as possible. That means shortcuts have to be taken. It is rare for an author to include all of the information she wants in a first edition of a book. Publishing new editions is expensive because publishers incur many of the same costs as a brand new book. That's why it takes at least 2-3 years before new editions are printed, but very few books are successful enough to warrant second or third editions.

Printed books, especially of the computer variety, contain incomplete information. While moving content online doesn't solve all schedule related problems, there doesn't have to be a definite end-date as there is with a book. Internet publishing allows for the content development process to be ongoing, which can result in more complete documentation over time.

### **3.1.6 Release content incrementally**

Most computer book publishers develop content in a waterfall-style model. It can take two years from the time an author writes the first word until the completed book makes it into stores. There is a significant opportunity cost between the time the content is ready for consumption and the time it is available to be consumed. Internet publishing allows for a more incremental model. Content can be released when it becomes ready.

This presents some problems to the traditional way of developing content. Generally, proof and copy editing occur after the final draft of a manuscript has been submitted by the author(s). To publish content incrementally would mean that copyediting would also need to happen incrementally. There is nothing inherently difficult about this other than aligning resources to do the work.

### **3.1.7 Update regularly and improve based on feedback**

In the print world, once a book is released, no major changes are made to it until a new edition is published. Minor changes can be made in between print runs, but nothing significant. Publishers can't revise new editions any quicker due to the production, inventory, and distribution costs.

For any given book, ten or fewer people read a significant portion of the manuscript before it hits the shelves. When the book becomes available to the public, readers can submit feedback to the publisher in the form of erratum, but this isn't very effective. Most customers aren't aware of the process of submitting errata and the ones that do may not feel it is worth the effort since the book won't be updated anytime soon.

With internet publishing, updating content is significantly less expensive. Publishers can actually take customer feedback seriously. Customers can provide feedback early and often, and the publisher can work with the author to get it integrated into the content. This is the first opportunity for publishers to create a true community around content unlike anything that was possible previously.

### **3.1.8 Extend the Long Tail**

Since writing a book is such a significant undertaking, many qualified people never attempt it. Publishers have to feel pretty confident that a book will turn a profit so they generally go after only mainstream technologies. There are a number of less popular technologies and products that are largely ignored by publishers.

Chris Anderson from Wired.com made the concept of the Long Tail famous in a 2005 article (see Figure 3 for a graphical summary) [12]. The figure shows how a significant portion of Amazon book sales comes from books that only sell a few units.

There is an even longer (invisible) tail of topics that publishers consider developing, but don't because they feel the demand isn't great enough to create a profit. Internet publishing drives all of the usual costs of publishing down such that those topics could be developed. With the Internet, it is not necessary to have a minimum of 10 chapters worth of content. Four or five chapters may be enough to pursue a topic.

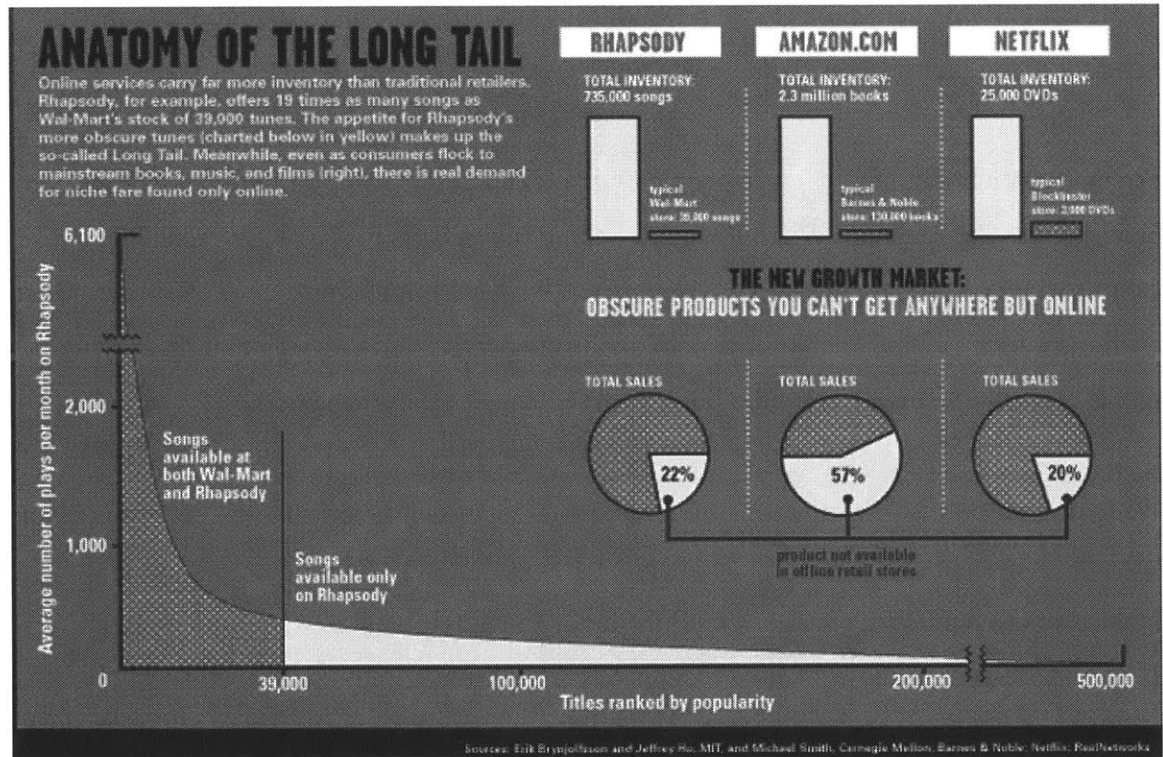


Figure 3. Anatomy of the Long Tail

### 3.1.9 Leverage the community

The transition from print to internet publishing is very disruptive. Publishers will be able to remix, customize, and do things with content never before imagined in the print world.

It is imperative to foster participation from the community of users to not only incorporate their knowledge in the content, but also stay in front of the disruptive trends.

One of the themes from the recent Web 2.0 phenomenon is the "architecture of participation" [13]. The idea behind the architecture of participation is that sites like Amazon, Craigslist, and eBay are successful in large part because they incorporate user-contributed content as part of their service.

As the unit of content decreases (from books to articles to even smaller pieces of information), the number of people that can contribute valuable content increases. The Internet allows publishers to pull from a larger author pool than print. The key challenge behind this will be separating noise from useful content.

### **3.1.10 Improve the author experience**

Authors are often treated like second class citizens in the computer book publishing world. Part of the problem is that publishers take on so much risk related to printing books that in return they get the upper hand when it comes to defining contract terms.

Most publishers aren't very good about keeping authors informed about their book(s). Some don't even provide timely sales data for an author's work. It is common for publishers to send quarterly or even bi-yearly royalty statements that have little useful sales data. Most publishers have computer systems that track timely sales data, but few have gone to the trouble of making this available to authors. Authors should be able to see at any time how well their content is selling and not be dependent on an ambiguous Amazon sales rank as their best measure for performance. The Internet provides a platform that can make creation of an author portal a worthwhile investment. Also, with Internet-based content, more information can be tracked in terms of who is accessing the content and from where. All of this could be provided to authors in an author portal.

## **3.2 Challenges with Internet Publishing**

While there are numerous advantages to using the Internet as a content delivery platform, there are a few serious challenges as well. Many of these challenges exist because internet publishing is a new form of publishing and there are few proven models. To top it off, the high rate of change with Internet technologies make internet publishing a constantly moving target.

To overcome these challenges, internet publishing firms have to do some experimentation and try different models to see what works and what doesn't. This in and of itself is problematic for the publishing world which has grown accustomed to a fixed and predictable way of doing business for decades. The publishers that are the most agile have the highest chance for success when using the Internet.

### **3.2.1 Creating profitable business models**

Coming up with sustainable and profitable business models is the number one issue publishers' face when considering internet publishing. This is a big unknown right now and it is causing many of the large publishers to wait on the sidelines until someone figures it out.

The Internet has fostered an environment where information is free flowing and openness is encouraged. As a result, you can find an enormous amount of quality content online for free. To date, attempts to monetize purely online content have had mixed results. There are several general models an internet publisher could employ which are explored in detail in Chapter 4.

### **3.2.2 Findability**

Let's assume a publisher has figured out a great business model and has created content customers are willing to pay for. The next challenge is making sure the target audience can actually find the content in the maze of information we know as the Internet.

An important player in this is obviously web search. However, if you have premium content that requires authentication, search engines won't be able to index the content. Not indexing "premium content" is a huge hole with web search today. Rumors last summer suggested that Yahoo and Google were working on this problem, but there have not been any recent updates. A new startup, Congoo, is trying to address the problem, but they require users to install a separate browser plug-in which will likely impede adoption.

One possible solution is referred to as cloaking. However, a recent story of Google pulling the German BMW web site from its index shows the negative consequences of cloaking [14]. In a nutshell, search engines don't condone the practice of displaying pages to users that are different than the pages it indexes (this is referred to as "cloaking"). In the BMW example, the page indexed by Google was full of keywords to help increase its Google PageRank while the page seen by users was much more scaled down and did not contain as many keywords.

Sidestepping the issue of Google acting as the sole judge and jury, the cloaking issue is a big deal for internet publishers. Google drives a huge amount of traffic on the web, but it can only do that for the content it indexes. In a subscription-based model, premium content has to be password protected which prevents search engine bots from indexing it. A publisher could have a lot of great content, but if it is password protected, search engines will never see it.

The alternative is to customize the response based on the requestor (i.e., cloaking). A publisher could give search engine bots unlimited access to the site while requiring authenticated access from users. However, if Google or any other search engines finds out about this, they may pull the content out of the index completely and without warning.

If cloaking isn't an option, what else is a publisher to do? For one, providing good descriptions on the free version of the premium pages can help so that most of the important keywords are captured by search engines. All the standard search engine optimization (SEO) techniques apply. However, even more important is to use other

means for driving traffic. Using free content (e.g., blogs) to drive traffic to paid content is a pretty common tactic today, although most publishers don't fully leverage this technique the way they could. Internet publishers have to find the right mix of paid and free content to attract customers.

### **3.2.3 Scaling the development of multimedia content**

As suggested earlier, internet publishers have an opportunity to use other forms of media such as audio, screencasting and web-enabled applications to enhance the learning experience. But it is not that easy. Video is considerably harder to author and edit than text. In fact, producing audio and video requires a completely different skill set. To make things worse, the skills that book authors and editors have aren't very transferable to audio and video. They are different skills.

One of the good things about the print business is that it is highly distributable. Authors don't need to sit beside editors to produce books. Multimedia content is trickier. It is important to have a good environment to record audio and video. The computer needs to be configured a certain way to make videos look consistent, and the room has to be conducive to recording or it will not sound good. Consistency in quality is something most publishers strive for and can obtain by making authors work within the confines of the publisher's environment (the templates, wordlists, etc.). It is not as easy to enforce restrictions in the multimedia world.

Then there is the issue that some of the best writers might not be that good at creating audio and video material. This presents a problem for publishers that want to encourage their existing author-base to work on multimedia content. One option could be to have writers focus on writing and have a separate group of people handle the conversion to audio/video, but that has its own challenges in coordination.



### **3.2.4 Usability on par with books**

One of the advantages of books is that they are very simple use. Online content sites have to strive for this same level of simplicity and usability.

Many content sites today are filled with obtrusive ads, some of which are positioned in such a way to trick users into clicking them. This is no way to establish a loyal customer base. When it comes to developing user interfaces for content sites, the requirements of users have to come first. Ads should be unobtrusive and context sensitive. In an ideal world, ads should enhance the user experience, but at a minimum they should not detract from it.

Another common mistake made by a lot of content sites is overloading their users. Some sites are so filled with links, panes, frames, sections, tabs, menus, and boxes that it is hard for users to find the content they came to the site for.

### **3.2.5 Fostering community involvement**

The social aspect of the Internet offers many prospects for publishers to incorporate feedback and original content from the community. This is new ground for publishers since most are not used to direct community involvement with publishing books.

Publishers have to become more social and reach out to the community much more than they have in the past.

There is a fine line between creating an environment where the community is active, involved, and contributes and one in which people feel publishers are just trying to take advantage of them. When O'Reilly first released their Roughcuts program which allowed people to buy electronic versions of books before they were ready for print, there was a small backlash [15]. O'Reilly encouraged readers to provide feedback on errors and problems they found in the text. Some people found this to be a veiled attempt to outsource quality assurance of O'Reilly books to the public.

### **3.2.6 Thinking beyond the book**

A challenge to computer book publishers as they migrate from print to the Internet will be thinking outside the boundaries of books. As discussed previously, a book is simply an organizational device that is necessary in the print world. Publishers and consumers alike have been thinking in terms of books for centuries. It will be difficult to overcome preconceived notions about how material should be organized, but unless we do, the full disruptive potential of internet publishing will not be realized.

### **3.2.7 Storing content in a portable format**

Most computer book publishers use Microsoft Word as the word processing tool of choice. While some publishers allow other tools such as OpenOffice, the costs associated with supporting each tool is quite high. Publishers use custom style templates that make conversion to a print ready format very easy. Each word processing application has its own way of creating and using templates, which makes the overhead associated with supporting each tool non-trivial.

The problem is slightly different when preparing content for the Internet. Publishers still need a mechanism to translate from word processor files, but the ultimate output is for the Web not print. The most flexible format to do this with is the extensible markup language (XML). Publishers will need to continue to support Microsoft Word and the other popular word processing applications, but the conversion process will consist of translating those files to XML.

Another option is to skip the use of traditional graphical word processing applications altogether and use a web application. O'Reilly has been working on a web-based environment to allow authors to develop and maintain books in the Hacks series.

### **3.3 Internet Publishing Models**

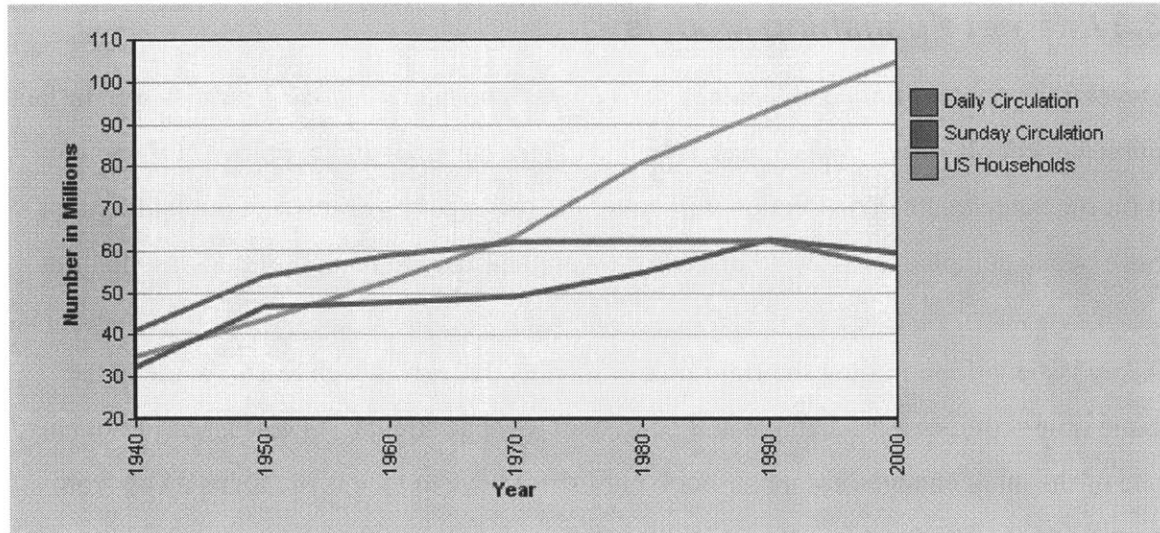
Developing content for the Internet is not a new phenomenon amongst publishers. In fact, publishers of all varieties have been trying to figure out what do about the Internet since it hit the mainstream. This section will cover the major forms of internet publishing that have been attempted to date and discuss their applicability to computer book publishing.

A key issue will be finding the right mix of models that drives web traffic while at the same time provides a predictable and consistent revenue stream. As discussed previously, one of the main changes computer book publishers have to face is the focus away from books and more toward free flowing content. Keeping an open mind about new possibilities is critical when evaluating internet publishing models because computer book publishers have the opportunity to enter new forms of publishing that weren't options to them in the print world.

#### **3.3.1 Online News**

The disruptive force of the Internet has been felt more in the news industry than perhaps any other area of publishing. Newspapers are fighting for their lives as consumers look to freely available news content online instead of paying newspaper subscriptions.

Newspaper circulation continues to decline year over year [16]. Figure 4 illustrates the point by comparing newspaper circulation with the number of households in the US since 1940. There is a noticeable decline that starts in the early 1990s that coincides with the introduction of the World Wide Web.



**Figure 4. U.S. Daily Newspaper Circulation versus Number of Households**

Source: Editor and Publisher Yearbook data; U.S. Census Bureau [16]

News content, the type that is short and fast changing, is the low hanging fruit of content on the Internet. Between free news sites such as CNN and MSNBC among many others and user-contributed content via blogs, most news stories are broken hours before newspapers go to print.

As far as computer book publishers go, technical news content can be a source to help establish a website as a useful destination for its readers. With today's technologies, it is easy for a publisher to create web pages with technology-specific news that contain filtered content for a specific area of interest (e.g., Microsoft or iPod). This can help with the "findability" problem discussed earlier as users find a publisher's web site through one means (news) and end up using it for another (premium content).

### 3.3.2 Online Magazines and Journals

Much as with the newspaper industry, print magazines and journals have been forced to move content online in order to stay relevant. However, their fate is also following that of the newspaper business. A recent article showed how two of the most popular Macintosh-related magazines, MacWorld and MacAddict, continue to suffer year over year losses in

the number of paid subscribers [17]. This story isn't unique as every year more magazines are forced to close due to lack of demand.

Most computer book publishers have not ventured into the magazine world, however, some have websites that act like online magazines. Sites like the O'Reilly Network provide article-length content on a variety of technical topics. O'Reilly doesn't charge users for this content, instead opting to include advertisements and promotional spots for relevant O'Reilly books.

Just as news content is fundamentally different than book content, so are articles. While authors can write articles as a basis for book content, articles are often more narrowly focused and tend to be more conversational. Just as O'Reilly has done, computer book publishers can use articles as a source of secondary revenue via ads, as a way to promote books, and another engine to drive traffic to the web site. Few computer-based magazines have been successful in charging for article-based content on the Internet and it is not likely computer book publishers would be able to either.

### **3.3.3 Web Books and E-Books**

In an ideal world for computer book publishers, they would be able to charge fees for electronic versions of their books that are comparable to the print versions and see comparable demand for both. So far, this has not happened.

Book publishers can publish book content online in a format such as PDF that makes the full book available in one file (what is generally referred to as an ebook) or integrated with the web as a series of HTML pages. There have been many attempts to make ebooks successful, but no solution has won out. The primary impediment to adoption is the lack of a ubiquitous ebook reader. Sony, Microsoft, and many others have attempted to create handheld devices that would replace the physical function of a book, but none have been widely accepted.

Ebook readers are crucial element of the ebook equation because most people do not like reading long PDF files on a computer. Computer screens do not provide the same quality or flexibility as the printed page. Even with that being the case, some computer book publishers have had success selling PDF versions of their books. Especially for books that are reference-oriented, PDF provides an efficient way to transport and search a book's contents. One problem with PDF is that piracy is a big concern to publishers. It is not uncommon for a book's PDF file to make it on an unprotected website on the Internet. Publishers have to keep constant watch for illegal distribution of their PDFs and send "cease and desist" letters when necessary.

Several vendors have created proprietary digital rights management (DRM) solutions to attempt to address this problem, but customers have not been receptive to buying PDF files that have built-in access restrictions. Some publishers, such as the Pragmatic Bookshelf, have opted for a simpler solution. Instead of requiring a heavy weight DRM solution, they simply put a footer on each page of the PDF with the name of the customer that ordered the book. That way, if one of their PDFs makes it on the Internet, they can easily identify who originally purchased it.

The other way to format book content online is with a web-based book. With a web book, the book's pages are converted to HTML. By far the leader in this space is Safari Books Online, which is a joint partnership between two of the largest computer book publishers: O'Reilly and Pearson. All O'Reilly and Pearson books are available on the web through Safari. Customers purchase "slots" which can be used to access a limited number of books per month [25]. Safari also offers corporate subscriptions, which make the entire library of O'Reilly and Pearson books available on the web for companies.

While Safari is a step in the right direction, it was just a mapping of the print business directly on the Internet. Safari doesn't provide any of the potential benefits for improving the development process as described in section 3.1. Nor does it facilitate community involvement or make use of multimedia content.

### 3.3.4 Blog Networks

The term "blog" was coined in 1998 as a derivation of "weblog", but online journals have been around since the early days of the web. In fact, blogging isn't as much of a technological innovation as it is a social phenomenon. It wasn't until blogging hit the mainstream that the web became a medium to capture society's stream of consciousness.

Blogs have been growing at an incredible rate. Technorati, a leading blog search engine has been keeping track of the growth in the blogosphere (the collection of blogs) since 2003. According to their stats, the number of blogs has doubled every five months for three straight years. Now they are tracking over 38 million blogs with 70,000 new blogs being added every day [18]. Figure 5 shows the growth curve.

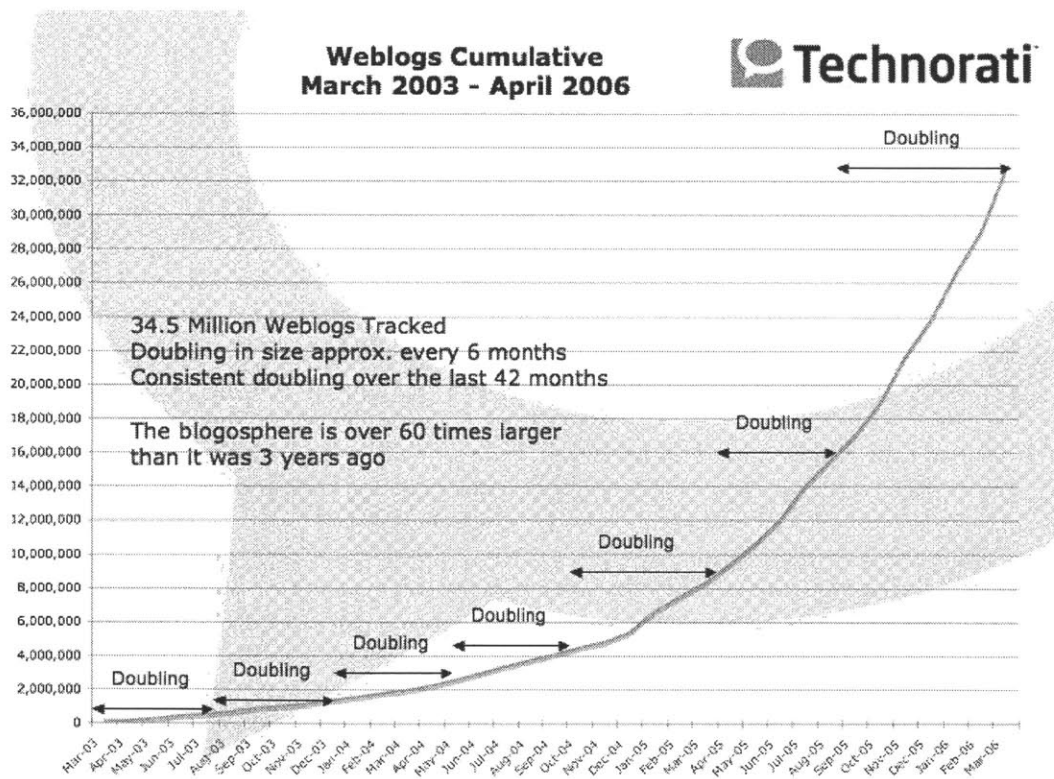


Figure 5. Cumulative Weblogs, 2003-2006

Blog authors monetize content through advertising. The more popular the blog, the more money can be made through advertising. Now, blogs have become a big business with a

couple of recent blog networks being acquired by large companies. America Online purchased Weblogs, Inc. in October 2005 for an estimated \$25 million [26]. Weblogs, Inc. is a network of over 85 blogs covering a variety of topic areas.

There are blogs for virtually any topic that can be imagined. Some computer book publishers have created blog portals for authors so instead of blogging on their own personal website, authors blog on the publisher's website. This is a great idea, but one that isn't being pushed as much as it could. The benefits of adding blog content to a computer book publisher's web site is similar to that of news or even articles. It provides another outlet for search engines to index a publisher's content and therefore another door a perspective customer may come through.

### **3.3.5 Content Marketplace**

Ecommerce marketplaces such as Ebay and Amazon have been extremely popular. On the other hand, there has yet to be a truly successful pure content marketplace. There have been several attempts, but no one has been able to create a place where customers are willing to pay for a broad range of online content.

The most recent example of a venture that attempted to create a content marketplace is the now defunct RedPaper.com. The purpose of RedPaper was to allow anyone to sell content for small amounts of money (i.e., under \$3). The combination micropayment and content marketplace did not attract enough revenue and despite much initial fanfare, RedPaper closed in 2005.

By combining several of the ideas discussed so far, such as creating content outside the boundaries of a book, content-as-a-service, and book content going online in the form of web books, one can imagine the need for a content marketplace for computer book publishers. The challenge, of course, is convincing customers of the value of purchasing small bits of content or even book-length content online. In an email exchange with RedPaper's founder, Mike Gaynor, he had this to say [8]:



"You had better offer content that is 10 X more compelling than anything they can find for free somewhere else on the web. You'll find that your potential revenue stream will look in 10 other places before they look in their wallet."

### **3.3.6 eLearning**

eLearning companies first started off by providing CD-based training. A customer could purchase CD sets that contained interactive material, tests, or videos to help learn a new technology or prepare for an IT certification. Now, some eLearning companies have web sites, such as CBT Nuggets (<http://www.cbtnuggets.com/>), where customers can access videos via the web. Many computer-based training (CBT) sites have been successful selling content in this model. In fact, based on the pricing for most CBT sites, eLearning companies are able to charge a premium for video-based content compared to what book publishers can charge for printed books.

### **3.3.7 Research Reports**

Research firms such as Jupiter, Gartner, and Forrester are worth mentioning because they've been able to monetize online content unlike anyone else. Recently, Jupiter Research released an electronic report on Google that cost \$750. The interesting part is the report was only 6 pages long. That comes out to \$125 per page [27]. A computer book publisher that produces a 400 page book that costs \$40 only gets roughly \$.05 per page (assuming a \$20 wholesale price). The obvious question that comes to mind is whether the research report is really 1000 times more valuable. Of course, the issue isn't as nearly black and white. Research firms typically charge outrageous fees for their reports so that their corporate customers maintain the perception that they are getting their money's worth. It also encourages potential customers to buy a corporate package instead of paying a high fee for a single report. Some may argue that content from analysts is more thoroughly researched and denser than a typical book, but that is debatable.

Based on the success of eLearning companies and Research firms at selling content online, in many cases which are comparable in quality to books and magazines, it illustrates how content is not created equal. Currently, you can charge a lot more for a book in a bookstore than you can for the same content as a PDF online. Similarly, you can charge significantly more for online content marketed as “eLearning” or “Research” than you can for books, news, articles, essays, and the like. How content is marketed plays a big role in how much a publisher can charge for it.

### ***3.4 Summary***

The Internet affords traditional computer book publishers the opportunity to correct many of the limitations that are inherent in the print industry. Because we are in the early stages of internet publishing, it will take some experimentation to figure out what works and what doesn't. The Internet also presents some problems to book publishers, the most significant of which is finding viable business models. This will be explored further in Chapter 4.

There are numerous publishing models that are available to book publishers, many of which offer options to explore areas that weren't available in the print world. The key is finding the right mix of models and the right mix of paid and free content to drive web traffic in addition to revenues.

## **4. Exploration of Business Models**

There are many different business models that a computer book publisher could use on the Internet. Finding the right mix of models is a critical first step for existing publishers as they consider the move online. This section will cover the standard methods for generating revenue online including the pros and cons of each model in the context of computer book publishing.

### **4.1 Free**

It doesn't make business sense for publishers to give away all their content (their core product) in return for nothing. This has worked in the software industry where some vendors, such as Red Hat, give away their software and make money through support and maintenance fees. In publishing this would translate into publishers giving away their content and upselling consulting services. However, given the independent nature of the publisher/author relationship, publishers do not have the staff to do consulting in a significant way without major organizational changes.

Giving content away for free isn't a viable business model by itself, but making some content free is a necessary evil. Many computer book publishers offer one or two sample chapters for download from each book for free. In addition to giving readers a taste of what a book is like, this helps address the findability problem discussed earlier as search engines index the content.

Searching premium or restricted content is not a feature of most search engines. The only way to drive traffic from search engines is through content that isn't restricted.

Fortunately, with the adoption of advertising as an accepted model on the Internet, publishers can make passive revenue with "free" content.

## **4.2 Advertising**

Virtually all internet publishers in existence today use some form of advertising as a source of revenue. It is very easy to implement and costs users nothing. Some researchers have argued that users incur a mental payment with advertising as they attempt to scan over ads [19], but ads are so common that most users consider it only a minor annoyance. Ultimately, advertising can be a good supplemental income stream, but in only a few situations (such as a high traffic blog network) can it be a viable standalone model for computer book publishers that have historically relied on royalties.

## **4.3 Subscription**

A very attractive model to publishers is subscription pricing. With subscription pricing, there is typically a tiered offering that allows customers partial or complete access to a full-range of content. This is easy to implement and provides a reoccurring revenue stream. In order to make this attractive to customers, new content must be continually added.

While this is a good model for publishers, subscription pricing tends not to be an attractive model for existing book authors. Subscription pricing works more like the magazine industry than the book industry. With magazines, authors are paid by the article. Regardless if one million or one hundred people buy the magazine and read the article, the author's income remains the same. With books, authors get an upfront advance, but get periodic royalties after the advance has been paid back based on how many copies are sold. That means for successful books, authors do much better than a publisher's best guess at what it will sell -- the advance. Subscription pricing levels the playing field. An author whose content is accessed once may be paid the same as one whose content is accessed hundreds of times.

Since delivery of the content would be driven through the Internet, it is possible to keep track of hits to each piece of content and pay based on frequency of use. However, this also has its downsides. If customers have open access to all content (i.e., site-wide

access), some may click through much of it while not really reading the material. In this case the number of hits doesn't truly reflect usage. There is also the potential to game the system whereby friends of an author access the author's content to artificially drive up the number of hits.

#### **4.4 Unit**

Paying per unit of content, such as a book or a magazine, is the standard print publishing business model. That translates on the web into paying for access to ebook versions of books. Publishers typically mark down ebooks by as much as 50% off the retail price of the print edition. The wholesale price of most computer books is around 50% of the cover price. When a publisher says that an ebook is marked down over 50% from retail, that doesn't mean the publisher is making any less from the book. In fact, publishers can make more from ebooks that sell at only a 50% discount from retail because there are no inventory and distribution costs.

O'Reilly has a Cookbook series that has been emulated across the computer book industry. An O'Reilly Cookbook contains a couple hundred "recipes" that show in discrete steps how to accomplish specific tasks or resolve problems with a particular technology. For task-oriented content, it would be possible to charge on a per-recipe basis or bundle some number of recipes together for a certain price. Customers could use "credits" to access the recipes of interests to them and purchase more credits when they run out. This would be equitable to authors because they could be paid based on usage unlike the site-wide access subscription model discussed earlier.

Pay per page and per chapter are other models that companies like Amazon and Google have considered [20]. There is quite a bit of skepticism about whether these models would actually work and no computer book publisher has announced plans to try something similar.

## **4.5 Sponsored**

A fairly recent strategy that has become more common is that of sponsored content. Sites like Realtimepublishers.com and more recently WindowsITPro.com have started publishing topical ebooks that are sponsored by one or more vendors. This is just another form of advertising as the vendors get their name associated with a book on a topic that is related to their business. The vendor can promote the ebook on their own website in an attempt to drive web traffic and demonstrate thought leadership for their area of focus.

Like any form of advertising, there is a catch. Most of the sponsored ebook sites require registration before a user can download the entire book. Part of the value proposition for vendors is that they can get demographic information about the users that download their ebook. On Realtimepublishers.com, users have to provide full name, postal address, and contact information along with occupation and field.

Print magazines have used a similar model for years. As the number of paid subscriptions decline, free subscriptions to people in positions of authority have become quite common among a variety of IT-related magazines. It remains to be seen if this will sustain magazines in the future or if it signals one last attempt to stay solvent.

Sponsored content is an area that book publishers have not even begun to tap in either the print or online markets. For books that sell really well, there is the possibility to have vendors sponsor specific chapters for a fee. In return, the vendor could be mentioned on the cover or within a chapter. The challenge for publishers that do this is to maintain the perception of neutrality and not seem overly biased in favor of the sponsors.

Another option is to write content about a vendor's applications or tools in exchange for a fee. For example, in a book about the Domain Name System (DNS), there could be a few chapters devoted to the DNS management tools of the leading vendors in the space. The benefit to the vendors would be to have their tools documented in a best selling book or widely read online content. The great thing about this model is that while it is a form of advertising, it is very targeted advertising. The audience is highly interested and looking

for solutions in the space. To make this attractive to vendors, the book or content would have to be one of the best sellers in the category or in high demand.

#### ***4.6 Earn***

A relatively untested model where users have to earn their access to future content plays to the social aspect of the Internet. With this model, a user would need to contribute, rate, tag, or review content in order to earn credits that can be applied toward the access of future content. This model can't work by itself because there is no revenue, but it could be a way to encourage participation from the community. Sites like Experts-Exchange.com have a credit-based model where "experts" accrue credits by answering questions on the site. They can use these credits to ask questions of other experts and review answers to questions.

#### ***4.7 Summary***

There are other variations on these models, but these are the major options in use today. Just as publishers need to use a combination of several internet publishing models (e.g., news and book content) it is also necessary to use a combination of business models. Publishers that can be nimble and adjust their internet business models as the market dictates will have the highest likelihood of success.

## **5. Improving the Content Development Process**

As mentioned in section 3.1.4, there are many similarities between writing code and writing content. There have been numerous books, articles, and even academic research on the topic of software development, yet relatively little attention is paid to the process of developing content. Since they have many commonalities, it is worth analyzing the ideas, processes, and procedures that can be borrowed from software engineering and applied to content, especially in the context of internet publishing.

### ***5.1 Comparison to Software Development***

First, it is useful to explore what specifically content development and software development have in common so we can focus our analysis on the areas that make sense. Much attention has been paid to software development processes because they can have a significant impact the quality and speed of delivery of a product. The same applies to content. One of the benefits of an author using a publisher versus self-publishing is the quality the publisher demands. Typically publishers have lengthy review and quality assurance cycles to ensure there are few errors in the end product. The key is figuring out how to optimize these processes.

#### **5.1.1 Similarities**

The act of writing has certain properties regardless of whether the output is intended for humans or machines. Writing books is about communicating to humans whereas writing software is about communicating with machines.

At the most basic level, the process used for developing software and content is the same. With both you design, develop, build, test, and support. The way each of these is done is drastically different, but the underlying process is the same.



The basic team structure for developing software and content is largely the same. Authors are like programmers. Editors are like project managers. Copyeditors and proofreaders are like QA testers. Since the underlying process and team structure are relatively the same for both, that means we should be able to use at least some of the processes from software engineering in regards to how a team functions and communicates.

Another similarity that this author has personally noticed working with numerous software developers and book authors is that both groups suffer from being overly optimistic when it comes to scheduling. Just as most programmers tend to overestimate their abilities to deliver code so do authors when it comes to estimating the delivery of chapters.

### **5.1.2 Differences**

Some aspects of developing software are fundamentally different from content. For example, software programs tend to be much more complex than even book-length content. A typical computer book will have anywhere from 100,000 to 150,000 **words** whereas it is not uncommon to have 100,000 **lines** of code in a commercial software application. Besides the sheer size, the magnitude of the complexity is also greater. It is relatively easy for a new author to come in and pick up the work from a previous author after reviewing the table of contents and a couple of chapters. It is difficult for a developer to come in midway through a project and become immediately productive as the learning curve to understand the whole software system is much greater than a book.

As the size and complexity of software projects increase so does the size of the team typically. A software development team could have anywhere from one to a hundred or more developers working at a time. When it comes to writing a book, there are rarely more than two or three authors on a single book. Programmers have to follow certain syntactical rules in order for their programs to compile correctly. With prose, there is no such restriction. That's why it is more difficult for multiple authors to work together; it is harder to have the content flow and read like it was written by one unified voice.

Over the years, the field of software development has matured to the point where many of the processes are heavily automated or assisted through tools. Integrated Development Environments (IDEs) are very robust and powerful, complete with syntax checking and correcting, built-in test harnesses, and even automated build process with a single command. Software developers have done a good job of using technology to help improve processes. Content development however is not much different than it was when Microsoft Word was first released. While features like spell and grammar checkers provide a definite benefit, authors do not have a full integrated set of tools that seamlessly ties all functions of the development process together.

Another area where software and content differ greatly deals with requirements and customer interaction. Section 2.2.5 covered the one-sided conversation publishers have had with customers. Lack of customer involvement has been a problem in the past with software too and new methodologies such as Agile stress the importance of getting customers involved early and often. Part of this translates into getting good requirements in which to build from. With content, authors don't work from requirements. There is no product requirements document (PRD) for a book. Instead, authors put together a table of contents and projected outline. Publishers will run this by noted experts in the field to get feedback. If no one raises any red flags about what's been proposed, the book moves forward with no customer involvement until the book is in stores.

A final difference worth mentioning is that of reuse. In the software world, reuse has been seen at times as a holy grail everyone should strive for. Why should developers code the same thing twice? Reuse is generally encouraged even though it often doesn't yield the efficiencies people hope for [21]. In content development, reuse is equivalent to plagiarism and is a very dirty word in the industry. That means each new book and article should contain original material. Some authors are able to work with a specific publisher and transition articles to book content and vice versa, but this is not the norm.

## **5.2 Applying Software Development Concepts**

Based on the analysis of the similarities and differences between software and content development, the four areas that will be explored further include the creation of content development methodologies, improving estimation techniques, putting more structure around the development process, and workflow automation.

### **5.2.1 Methodology**

One of the first questions any programmer will ask when starting a new project is: What methodology will we use? Early on, those in software engineering identified a need for a well constructed set of principles that guide the flow of development and identify key priorities that should not be overlooked. A methodology brings some amount of predictability to an otherwise chaotic activity. There are over a dozen development methodologies with the most popular including Waterfall, Agile, XP, Iterative, Spiral, and RUP.

Interestingly enough, people in publishing never talk in terms of development methodologies. The book development process is fairly loose with minimal guidelines and rules. Implementing a methodology has many benefits that could benefit publishers in the following ways:

- Reduce the learning curve for new authors by using a standard set of processes
- Clearly define expectations among all parties
- Provide focus on what's important
- Establish a repeatable process

As mentioned in Chapter 2, if one had to pick a methodology that most closely matched how computer book publishers develop content, it would be Waterfall. The Waterfall methodology is considered a predictive method categorized by a series of stages that are highly planned at the outset of the project. There have been many criticisms against Waterfall including its inflexibility to change, the requirement to design everything up

front, late stage integration, and disconnect between designers, developers, and customers [28]. Many of these problems also exist with the way book content is developed today.

A more recent methodology called Agile has achieved a lot of popularity and has applicability to internet publishing. Agile was created in 2001 by a group of 17 developers that represented various methodologies. The result was the Agile Manifesto [22]. Reproduced below are the Principles behind the Agile Manifesto:

Our highest priority is to satisfy the customer through early and continuous delivery of valuable software.

Welcome changing requirements, even late in development. Agile processes harness change for the customer's competitive advantage.

Deliver working software frequently, from a couple of weeks to a couple of months, with a preference to the shorter timescale.

Business people and developers must work together daily throughout the project.

Build projects around motivated individuals. Give them the environment and support they need, and trust them to get the job done.

The most efficient and effective method of conveying information to and within a development team is face-to-face conversation.

Working software is the primary measure of progress.

Agile processes promote sustainable development. The sponsors, developers, and users should be able to maintain a constant pace indefinitely.

Continuous attention to technical excellence and good design enhances agility.

Simplicity--the art of maximizing the amount of work not done--is essential.

The best architectures, requirements, and designs emerge from self-organizing teams.

At regular intervals, the team reflects on how to become more effective, then tunes and adjusts its behavior accordingly.

Nearly all of this can apply to content development just as easy as it does to software development--simply replace the word "software" with "content" above. In fact, an Agile style closely relates with many of the opportunities to improve the content development process as outlined in section 3.1.

There are other Agile-related methodologies that have applicability to internet publishing. Iterative methodology suggests developing software in an incremental or iterative style where learnings are applied throughout the lifespan of the software. Extreme Programming encourages continuous integration and small releases to ensure there is always a working system. This is a useful mindset in the context of content because it gets authors used to constantly releasing bits of content instead of waiting until the end of the project to put all the chapters together.

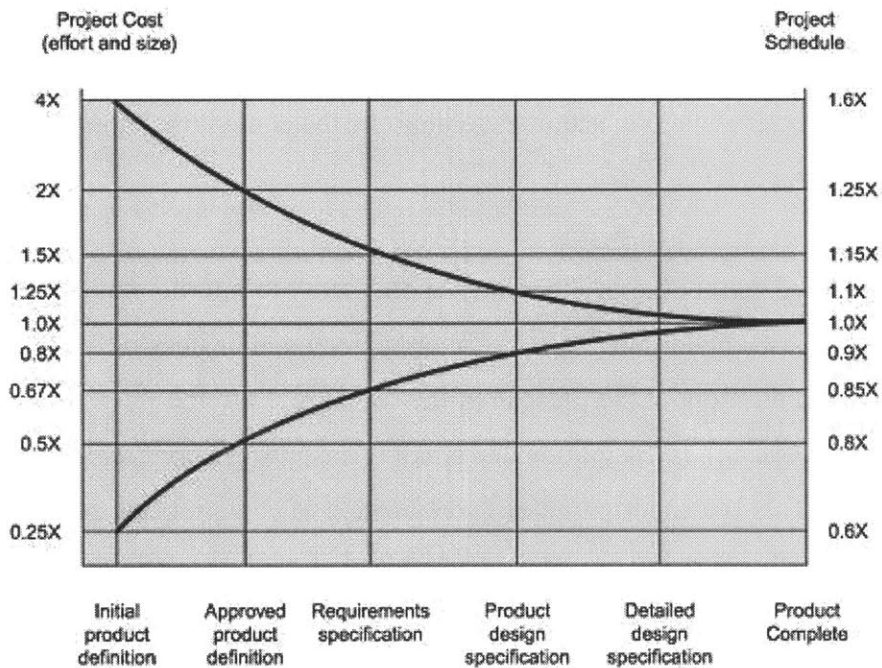
The point here is not that a single software development methodology neatly fits content development, but that much could be taken from existing methodologies to create a holistic system for content development.

## **5.2.2 Estimation**

Much like software developers, book authors are notoriously bad at estimating how long it will take to complete a book. When it comes to computer books, the vast majority of authors are writing as a second job or hobby; it is rarely the primary source of income. As

a result, authors typically juggle multiple projects at the same time and often have time for writing only at nights or on the weekends. This makes the job of estimating even more difficult.

Steve McConnell's Rapid Development [29] has a wealth of information on the software engineering process and has a whole chapter dedicated to estimation. The graph in Figure 6 from McConnell's book is a widely used illustration to show the impact of estimation on project cost over the life of a project.



**Figure 6. Estimate-Convergence Graph**

Source: Rapid Development (1996)

In the early stages of a book, the precision of the schedule is greatly variable, but improves as content gets developed.

Software estimation is extremely hard because it is difficult to nail down exactly what the end-product is suppose to be, not to mention what it will take to develop. Content development is a little more straightforward. It is easier to scope the entire project.

Consequently, if publishers spent a little more time analyzing the estimation problem, they are likely to see very positive results.

McConnell suggests a three step process to create an accurate development schedule:

1. **Estimate the size of the product.** In this case the product is either a book or content that will be made available on the Internet.
2. **Estimate the effort.** This is done in man-months.
3. **Estimate the schedule.** Based on the size and effort, the number of calendar months can be determined.

The size of a writing project can often be characterized by the number of words or chapters. This often is a good measure, but needs to be tweaked based on the type of book. For books that contain source code, the code needs to be counted differently, often with a lower weight than words. Also, some content may be very graphical and contain fewer words, which would need to be accounted for. Lastly, projects that consists of screencasting have to be measured in a completely different fashion since there aren't written words to count. An area to research further would be to create a function-point system similar to what McConnell proposes, except based on the intricacies of developing content.

The other important element of the schedule estimation equation is effort. This is something that publishers can collect a plethora of data on, but don't currently. If publishers did a better job of recording the amount of work authors do and how long it takes them to do it, undoubtedly a curve could be mapped to the data.

The last step is computing the schedule estimate. McConnell's equation is the following:

$$\text{schedule in months} = 3.0 \times \text{man-months}^{1/3}$$

Is this right equation for content? Some question whether it is right for software. Since this is uncharted territory in the content world, it will require some experimentation and tweaking of the equation to find a fit.

This is just one method to create an estimate. Publishers should start tracking more metrics around author estimates and their accuracy so that the art of estimation can become more predictable in the content world.

### **5.2.3 More structured development process**

With the Waterfall style development that most books follow, having a structured development process hasn't really been necessary. The integration of chapters wasn't needed until the very end because copyeditors typically don't start until the final author draft is completed. However, with internet publishing there are several reasons to create more structure around development similar to what is done for software.

If internet publishing allows for content to be distributed as soon as it is ready, that means the underlying production processes need to be able to handle that. With books, there is no or very primitive revision control enforced by the publisher. With a rapid deployment scenario, it will be necessary to have an environment where changes can be made quickly and tracked over time. This calls for some sort of revision control system where authors check content in and out much like software developers do for code.

Likewise, it will be necessary to have a build process so that content can easily be transitioned from the format the author provides to the Internet. Today, the conversion from word processing files to a print-ready format is tedious and often manual. This will not scale in an internet publishing model that requires the ability to update content continuously. Also, having an automated build environment could put more responsibility for following template guidelines back on the author. Typically, a production editor is responsible for sifting through an author's work and manually fixing any discrepancies. With an automated build environment, some level of automated testing for things like



style and format could be incorporated. A final benefit of having a robust build process that is the ability to customize the content based on the customer. The Pragmatic Bookshelf used such a system to provide a cheap DRM solution by automatically adding the customer's name to the bottom of every page of a book's PDF file.

Having a central repository for content provides many other benefits including the option to create a workflow system.

#### **5.2.4 Workflow automation**

As mentioned in section 2.2.4, a workflow system could automate many of the manual communications that go on between authors, editors, and tech reviewers. Basic workflow systems are commonplace in software development shops. Developers are automatically notified when they need to take action, for example, when bugs are submitted for their code, when they are assigned new action items, or when unit tests fail for their code.

The same concept can be applied to content development. Besides the benefit of increased communication, publishers would be able to instantly get a status of a content development project. Much of the time that is wasted today in terms of an editor forgetting to contact a tech reviewer or a tech reviewer forgetting when a chapter was due could be alleviated with a fairly simple workflow system.

### **5.3 Summary**

This chapter is just the tip of the iceberg as far as possibilities for formalizing the content development process. With all of the research and years of learning that have occurred in the field of software engineering, much can be borrowed to establish a basis for content development methodologies, processes, and ideas.

## **6. Conclusions**

The main goal of this thesis was to analyze the computer book industry and identify the opportunities and challenges with using the Internet as a content delivery platform. The primary conclusion drawn from the research and analysis is that traditional publishers are putting their businesses at risk unless they more actively explore Internet-enabled publishing methods and business models to capture a portion of the increasing demand for online content. There are difficulties that have not been completely worked out, but the current trends with print-based publishing show the consumer migration away from print especially in the areas of news and computer reference information. Publishers that remain devoted to print will continue to see their businesses decline.

### ***6.1 Important findings***

The following list highlights the important findings from this thesis:

- Computer book sales map closely to overall IT spending and are still down more than 40% from the 2001 highs. The quality of free documentation on the Internet is starting to rival the quality of books, which is also eating into the computer book market. Publishers have an opportunity to extend the "Long Tail" by developing content for smaller niche topics, which may be a source of significant revenue.
- A disproportionate share of the top forty computer books has gone to books that are more readable and entertaining than the typical computer book. Computer book publishers need to focus more heavily on the entertainment value of content.
- There is a significant opportunity cost during the development of a book as content written first must lie dormant until the rest of the book is complete. The Internet allows for a more incremental and iterative publishing model.
- The concept of a book is largely an artifact of the print world. The Internet provides an opportunity to think about organizing content in different ways, but

this requires both publishers and consumers to think outside the boundaries of books.

- Internet publishers will not have the same level of control in the industry as traditional book publishers. Instead of being a necessary source of capital to fund content development, internet publishers will have to create value-added services to attract and retain authors.
- How content is marketed plays a big role in how much customers are willing to pay (if at all). eLearning and Research web sites are able to charge a premium for content whereas book, news, and magazine publishers have had a harder time.
- Creation and adoption of a content development methodology has many potential benefits for publishers. Internet publishing requires adoption of a more structured development environment to allow for quick turnaround of content. A workflow system could drive out many of the inefficiencies in communication and coordination that unnecessarily delay book projects.
- To be successful, computer book publishers will need to mix and match internet publishing models and business models to find the right combination based on business objectives. This will include using a combination of advertising, sponsored content, and pay per unit. Publishers need to be nimble due to the high degree of experimentation that is required to find a working model. This will most likely require organizational changes as most publishers do not have a culture of innovation.

## **6.2 Future research**

This thesis has been restricted to a small subset of the overall book publishing industry, namely computer book publishing. While many of the issues covered here are applicable to other forms of publishing, the nuances are different enough so that broad generalizations cannot be made across the industry without further research. It is possible, if not probable, that additional research on the industry at large would yield similar findings to this thesis. For example, many of the benefits of internet publishing covered in section 3.1 are relevant to non-fiction and fiction publishing. Additionally, many of the

issues covered in chapters 4 and 5 related to business models and the content development process are relevant to all forms of book publishing. Broader research across the book publishing industry may identify other opportunities not immediately obvious from research of the computer book publishing niche.

As mentioned in Chapter 5, the content development process has been largely ignored and could benefit from significantly more study and research. Key issues were highlighted in Chapter 5 that could enable computer book publishers to improve the efficiency and predictability of their business along with setting them up to take full advantage of the Internet through better development processes and practices.

## Bibliography

1. Book. (2006). *Encyclopædia Britannica*. Retrieved April 4, 2006, from Encyclopædia Britannica Premium Service.
2. Publishing, History of. (2006). *Encyclopædia Britannica*. Retrieved April 6, 2006, from Encyclopædia Britannica Premium Service.
3. The US Book Publishing Industry Consists of About 2,600 Companies with Combined Annual Revenue of \$30 Billion. (2006). Retrieved March 30, 2006 from <http://biz.yahoo.com/bw/060330/20060330005306.html?.v=1>.
4. Poynter, Dan. (2004). *Statistics*. Retrieved January 10, 2006 from Para Publishing site: <http://www.parapublishing.com/sites/para/resources/statistics.cfm>.
5. About Nielsen BookScan. (n.d.) Retrieved February 16, 2006 from <http://www.bookscan.com/about.html>.
6. O'Reilly, Tim. (2005, April). *Book Sales as a Technology Trend Indicator*. Retrieved January 12, 2006 from O'Reilly Radar site: [http://radar.oreilly.com/archives/2005/04/book\\_sales\\_as\\_a.html](http://radar.oreilly.com/archives/2005/04/book_sales_as_a.html).
7. O'Reilly, Tim. (2006, April). *State of the Computer Book Market, Part 1*. Retrieved April 22, 2006 from O'Reilly Radar site: [http://radar.oreilly.com/archives/2006/04/state\\_of\\_the\\_computer\\_book\\_mar.html](http://radar.oreilly.com/archives/2006/04/state_of_the_computer_book_mar.html).
8. Mike Gaynor (personal communication, March 8, 2006) discussed the collapse RedPaper.

9. Waterfall Model. (2006). *Wikipedia*. Retrieved April 5, 2006, from [http://en.wikipedia.org/wiki/Waterfall\\_model](http://en.wikipedia.org/wiki/Waterfall_model).
10. D. Thomas in a presentation at O'Reilly Foo Camp 2005.
11. Wakefield, Maureen. (2003). *Lean-In Technology, Safety Nets and Psychology*. Retrieved January 12, 2006 from ITForum site: <http://it.coe.uga.edu/itforum/paper67/paper67.htm>.
12. Anderson, Chris. (2004, October). The Long tail [Electronic version]. *Wired Magazine*, 12.10.
13. O'Reilly, Tim. (2005, September 30). *What is Web 2.0*. Retrieved April 22, 2006 from O'Reilly Network site: <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>.
14. BMW given Google 'death penalty'. (2006, February 6). *BBC News*. Retrieved March 2, 2006 from <http://news.bbc.co.uk/1/hi/technology/4685750.stm>.
15. Lim, Kevin. (2006, January 25). *O'Reilly's RoughCuts: Shouldn't they pay you instead?*. Retrieved January 30, 2006 from <http://theory.isthereason.com/?p=676>.
16. Journalism.org. (2004). The State of the News Media 2004. *Annual Report on American Journalism*. Retrieved April 24, 2006 from [http://www.stateofthenewsmedia.org/chartland.asp?id=174&ct=line&dir=&sort=&col1\\_box=1&col2\\_box=1&col3\\_box=1](http://www.stateofthenewsmedia.org/chartland.asp?id=174&ct=line&dir=&sort=&col1_box=1&col2_box=1&col3_box=1).
17. U.S. Mac magazines suffering reader drought. (2006, April 17). *Think Secret*. Retrieved April 29, 2006 from <http://www.thinksecret.com/news/0604magazines.html>.

18. Sifry, Dave. (2006, May 1). *State of the Blogosphere, April 2006 Part 2: On Language and Tagging*. Retrieved May 2, 2006 from Technorati Blog: <http://technorati.com/weblog/2006/05/100.html>.
19. Everard, Andrea, Scott McCoy, Dennis Galletta, and Peter Polak. (2004, September 9). *A Study of the Effects of Online Advertising: A Focus on Pop-Up and In-Line Ads*. Retrieved April 25, 2006 from <http://mis.temple.edu/research/documents/GallettaPopupAdStudy.pdf>.
20. Amazon to offer pay-per-page e-books. (2005, November 4). *Vnunet*. Retrieved January 29, 2006 from <http://www.vnunet.com/vnunet/news/2145517/amazon-random-house-launch-pay>.
21. Schmidt, Douglas. (1999, January). *Why Software Reuse has Failed and How to Make It Work for You*. Retrieved April 28, 2006 from <http://www.cs.wustl.edu/~schmidt/reuse-lessons.html>.
22. Principles behind the Agile Manifesto. (n.d.) Retrieved February 16, 2006 from <http://agilemanifesto.org/principles.html>.
23. Locke, Chris, Doc Searls, and Dave Weinberger. (2001). *Cluetrain Manifesto*. 1st ed. Perseus Books Group.
24. Pike, Robert W. (2003). *Creative Training Techniques Handbook*. 3rd ed. Amherst: Human Resource Development Press.
25. Safari Products. (n.d.) Retrieved April 30, 2006 from <http://www.safaribooksonline.com/products/overview.asp>.
26. Ali, Rafat. (2005, October 5). *Exclusive: Weblogs Inc Being Bought Out By America Online*. Retrieved January 25, 2006 from

[http://www.paidcontent.org/pc/arch/2005\\_10\\_05.shtml#051691](http://www.paidcontent.org/pc/arch/2005_10_05.shtml#051691).

27. Allen, Robbie. (2006, March 24). *JupiterResearch Google Report: \$125 per page*. Retrieved May 1, 2006 from <http://netawaremedia.com/blog/2006/03/24/jupiterresearch-google-report-125-per-page/>.
28. Criticism of the Waterfall Model. (2006). *Wikipedia*. Retrieved May 1, 2006, from [http://en.wikipedia.org/wiki/Waterfall\\_model#Criticism\\_of\\_the\\_waterfall\\_model](http://en.wikipedia.org/wiki/Waterfall_model#Criticism_of_the_waterfall_model)
29. McConnell, Steve. (1996). *Rapid Development*. 1st ed. (163-204). Redmond: Microsoft Press.