

GOING DIGITAL: A GUIDE FOR BOOK PUBLISHERS

by

MariÈve Mauve Pagé

Bachelor of Arts (Honours), Queen's University (2002)

Bachelor of Communication Design (Honours), NSCAD University (2005)

PROJECT SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Publishing

In the Faculty of Arts and Social Sciences

© MariÈve Mauve Pagé, 2008

Simon Fraser University

Spring 2008

All rights reserved. This work may not be reproduced in whole or in part,

by photocopy or other means, without permission of the author.

A P P R O V A L

N A M E: MariÈve Mauve Page

D E G R E E: Master of Publishing

T I T L E O F P R O J E C T: Going Digital: A Guide for Book Publishers

S U P E R V I S O R Y

C O M M I T T E E: _____

Dr. John Maxwell

Senior Supervisor

Assistant Professor, Master of Publishing, Simon Fraser University

Dr. Rowland Lorimer

Supervisor

Professor and Director, Master of Publishing, Simon Fraser University

Karen Gilmore

Supervisor

Vice-President Operations, Corporate and Legal Affairs,
Douglas & McIntyre Ltd.

D A T E A P P R O V E D:

January 28, 2008



SIMON FRASER UNIVERSITY
LIBRARY

Declaration of Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the "Institutional Repository" link of the SFU Library website <www.lib.sfu.ca> at: <<http://ir.lib.sfu.ca/handle/1892/112>>) and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author's written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada

A B S T R A C T

This project report, structured as a guide, strives to inspire and assist small-to-mid-sized Canadian trade publishers to develop their digital strategies. The need for digitization in a period of transition within the publishing industry is explored, as well as the different steps to be taken to create a successful digital strategy. This guide first explores the goals and motivations of digitization, specifically looking at websites, viral marketing, book browsing and searching, and e-books. It then reviews the types of rights necessary for a digital strategy, and the decision making process necessary for the selection of titles to be digitized. Finally, the guide explores the different formats and platforms available for digitization and looks at current efforts to standardize them.

Keywords: Digital strategy; Digitization – Books; Digital rights; Digital assets; Online marketing; Trade publishing

Subject words: Digitization; Digital rights; E-books; Online marketing; Trade publishing

A C K N O W L E D G E M E N T S

My sincere thanks to my supervisor John Maxwell for his support and constructive comments. I would also like to thank all the staff members at Douglas & McIntyre, specifically Karen Gilmore, Allison Urowitz, Emiko Morita, and Rob Sanders; without their help and time this project could not have come to completion.

Finally, I wish to thank Isabelle Boisclair for her encouragement and comfort, and all my classmates and professors at Simon Fraser University for a great experience in the Master of Publishing program.

C O N T E N T S

ii	APPROVAL
iii	ABSTRACT
iv	ACKNOWLEDGEMENTS
v	CONTENTS
vii	LIST OF TABLES
viii	GLOSSARY
1	INTRODUCTION
1	Douglas & McIntyre
3	The guide
6	CONTEXT
6	Introduction
6	Changing times
8	Possibilities
10	Constraints
11	The unknown
12	Free versus paid content
18	Section conclusion
19	WHY GO DIGITAL? (ADVANTAGES AND MOTIVATIONS)
19	Introduction
19	Websites strategies
20	Publishers' websites
22	Websites created for authors and books
22	Blogs and social networks
27	Viral marketing
30	Book browsing and search engines
30	Book browsing
32	Search engines
32	Google
34	E-books
37	Section conclusion

38	WHAT DO YOU NEED? (DIGITAL RIGHTS AND TITLES)
38	Introduction
38	Types of rights
39	Primary rights
40	<i>Formats</i>
40	<i>Transmission/distribution</i>
40	Subsidiary rights
41	Marketing rights
42	Competing works
42	Ownership and procurement of rights
43	Choosing titles to be digitized
48	Section conclusion
49	WHAT DOES DIGITIZATION LOOK LIKE?
	(FORMATS, PROCESSES AND MANAGEMENT OF DIGITIZATION)
49	Introduction
49	Formats
49	Flowable text
50	Paginated text
51	Processes to create digital content
52	Image scanning
52	OCR (Optical Character Recognition)
53	Tagging
53	Metadata
53	Managing and distributing digital assets
53	DAMs and DADs
55	In-house versus third-party service providers
59	Standardization
60	Section conclusion
62	CONCLUSION
65	BIBLIOGRAPHY
72	APPENDICES
72	Appendix A: E-book features
73	Appendix B: Creative Commons
74	Appendix C: DAM(D)s and conversion pricing

L I S T O F T A B L E S

p.25 **Table 1:** Examples of digital formats and the type of material that can go online

G L O S S A R Y

Blogger

A blogger refers to the creator of a blog (or “web log”), who adds content to a website that is displayed chronologically. Bloggers write anything from personal journal entries to comments on specific and specialized subjects. Popular bloggers can generate income from their online activity, primarily from advertising revenue.

Digital Asset Distributor (DAD)

DADs are service providers that distribute the digital content (or digital assets) of a publisher to e-book distributors and retailers. Most DADs also help publishers with the preparation of digital files and the process of digitization.

Digital Asset Management (DAM)

DAMS are service providers that manage and maintain the publisher’s digital assets.

Digital Asset Management and Distribution (DAMD)

DAMD refers to the combination of the service provided by DAMs and DADs. As combined service providers, most DAMDs are retained on contract to implement the technical aspects of a publisher’s digital strategy, which implies producing, managing, storing, and distributing the digital assets.

Digital Rights Management (DRM)

DRM is digital technology that monitors the use of digital files and can manage and prevent access and use of the file. DRM is used to protect a work’s copyright.

Digital strategy

A publisher's digital strategy is the plan that analyzes and outlines, for an individual title or an entire list of titles, the different needs for digital content (such as digital versions of a title or multimedia digital content, such as videos and photographs) and outlines the digitization process. The digital strategy aims to benefit from the publisher's digital assets, by reaching a greater reader audience and maximizing profit.

Digitization

Digitization is the process of making a title available in any digital form (which can include making the entire book available digitally or creating additional digital content to accompany the title, such as enhance or multimedia digital content).

eXtensible HyperText Markup Language (XHTML)

XHTML is a markup language based on HTML but realized as an XML Document Type.

eXtensible Markup Language (XML)

XML is a markup language where users create their own tags and attributes to meet the needs of the document.

GIF

GIF (*Graphics Interchange Format*) is an image file format commonly used on the web. GIFs have a bit depth of 8-bit per pixel (i.e. images are made of a combination from 256 colours). GIFs are useful for illustrations viewed online.

HTML (Hypertext Markup Language)

HTML is a markup language used on the web; it defines the structure of the text by identifying elements such as the format and the presentation of the document.

JPEG

JPEG (*Joint Photographic Experts Group*) is a highly-compressed image file format. Compared to GIFs, JPEGs do not have a limited amount of colours and are useful for photographic images viewed online.

Metadata

Metadata is information about the data or content. This information is used to better manage and use the content. For example, metadata of a book title can be the ISBN number, the publication date, the dimensions of the book, and the names of the author(s) and contributor(s).

Open Packaging Format (OPF) Specification

OPF Specifications refer to standards creators of e-readers must follow to ensure that EPUB files can be read on e-readers. The standards are developed by the International Digital Publishing Forum (IDPF).

Open Publication Structure (OPS) Specification

OPS specifications describe the standards for the content of e-book file format EPUB. They are guidelines to ensure consistency of EPUB files from content creators. The standards are also developed by the International Digital Publishing Forum (IDPF).

Optical Character Recognition (OCR)

OCR is a mechanical process using software that converts a scanned image of a page to machine-readable text.

Personal Digital Assistant (PDA)

PDAs are hand-held devices that provide access to the Internet and most often include a cell phone, providing users with the capacity to email and search the web. PDAs are essentially small computers that users can use as word processors, calculators, video game devices, music and video players, and e-readers.

Print-on-demand (POD)

Print-on-demand is a digital printing method where a copy of a work is created only after the printing sale order is received. POD's impact on the book publishing industry is noticeable because of short print runs: for a small amount of copies, POD is less expensive and faster than offset printing.

Social networks (online)

Online social networks are online communities created by social network services, such as MySpace and Facebook, that provide communication services such as instant messaging, file sharing, blogging, discussion spaces and forums. Social networks are often centered on people's interests or activities. They are typical of "Web 2.0."

Viral marketing

Viral marketing is marketing online that depends on web users to share and distribute content to create online brand awareness and word-of-mouth.

Web 2.0

Web 2.0 refers to the so-called second generation of the Web. Unlike the first “version,” Web 2.0 is known for its user interactivity. Web 2.0 is based on online communities that interact with each other and with the content that populates the web (by sharing, manipulating, and developing content).

Widget

Widgets are web application code embeddable on other HTML-based websites. They are similar to a small website placed on a bigger one. Book widgets refer to a widget produced and distributed by a publisher that can be shared and embedded on third-party websites, such as bloggers’ and retailers’ sites.

I N T R O D U C T I O N

After a first wave of e-books a few years ago,¹ the trade publishing industry is seeing once again talk of e-books and e-readers. However, this time the situation is different. In the last few years, trade publishers have started to follow the lead of scholarly publishers, and have begun to analyze the possibilities of digital content that goes far beyond e-books. The digital market has become more accessible in the last few months with many conferences and forums exploring the possibilities of broadening the digital scope for trade publishers (conferences by groups such as BookNet Canada, Book Expo America, and O'Reilly).

D O U G L A S & M C I N T Y R E

In the course of an internship at the Canadian publisher Douglas & McIntyre, I had the privilege to witness the early development of the company's digital strategy. This project report is informed by Douglas & McIntyre's process in developing this strategy. The report aims to guide and inspire other publishers in similar processes. The focus of this report is specifically on small-to-mid-sized Canadian publishers as they develop their own digital strategies. This guide strives to keep in mind that individual publishers have different and specific needs, activities, and title lists, and to advise them through the transitional period faced by today's publishing industry. The publishing industry no longer relies only on the printed word, but rather is moving towards incorporating digital technologies in all aspects of the publishing process. From conception to delivery, digital technologies are part of publishing: files are transferred via the web; design and production

1 The first e-books, as we know them, were available in the late 1990s. But the possibility for e-books has been present since books were layed out digitally; it is only with the popularization of the Internet that e-books became a financial possibility. However, the historical origins of e-books date from 1971. In fact, Michael Hart started what became the Project Gutenberg (www.gutenberg.com) with the creation of "Etext"; Hart believed in the potential of storing, retrieving and searching content through digital libraries (www.gutenberg.org).

departments rely on digital files; and more importantly, book marketing and promotional efforts are increasingly dependent on the web, through technologies such as search engines, viral marketing, and blogs.

Like many other trade publishers, Douglas & McIntyre initially had some concerns regarding the digitization of their titles. The major obstacles for Douglas & McIntyre to digitize titles were the unanswered questions about that process:

- How does a publisher manage and organize digital files and digital rights?
- How does one deal with copyright issues for digital formats and how do the rights vary for full versions of the work, partial content made available digitally, and enhanced digital content?
- How does a publisher deal with author contracts for books dating *before* the development of digital technologies such as e-books, the Internet, and Personal Digital Assistants (PDAs)?

To address these concerns, Douglas & McIntyre decided to take part in various controlled and limited projects. Douglas & McIntyre's first digital project was to join the Amazon "Search Inside the Book" program in 2005. This project was of a limited scope, and the terms of the project were dictated by Amazon. Amazon scanned publishers' titles for free and included them in the "Search Inside the Book" program. Unfortunately, publishers never had access to or control over the digital files produced by Amazon. This program was a first step toward digitization, but was too limited for publishers who aimed to fully explore the possibilities of digitization.

For Douglas & McIntyre, the real beginning of their digital strategy came with the Best of BC Digital Reference Collection,² a pilot project started by Gibson Library Connection Inc. and the Association of Book Publishers of British Columbia (ABPBC). The Best of BC Digital Reference Collection is a collection of books in digital format from BC publishers that will be submitted to the Canadian Electronic Library (CEL) to be made available to three key library consortia in

2 Formerly known as the Best of BC Digital Library Collection.

British Columbia.³ This project was an opportunity for Douglas & McIntyre to test the possibilities associated with the digitization of books, and a chance to analyze and understand the possible digital markets and different uses of digital files, all within a controlled environment. The project had specific guidelines regarding the subject and the number of titles selected, and had a distinct established audience. The books submitted to CEL for the Best of BC Digital Reference Collection will be accessible through university, college, school, and public libraries, making researchers and students its primary audience.⁴

THE GUIDE

This guide describes the steps taken by Douglas & McIntyre to begin the digitization process, and aims to provide valuable information to small-to-mid-sized Canadian publishers in the process of developing their own digital strategies. This guide answers common questions publishers have about the digitization process. The goal is *not* to speculate about an undefined and continuously changing digital future, but rather to find solutions to current issues and questions. The trade publishing industry is in transition; this guide is meant to help publishers become more flexible during that period. This guide also demonstrates that the digitization process affects all departments within the publishing house. Digitization affects a title at its early conceptual stage, because decisions need to be made regarding the online marketing of a title and the likelihood the subject can translate well unto the online sphere. Digitization also affects the editorial department since they are most often the ones responsible to manage and enter metadata. Digitization has an impact on the procurement of rights by influencing the types of digital rights necessary for the different aspects of the digital strategy. Digitization influences potential marketing decisions and promotional opportunities, and the production and design processes. Finally, digitization affects

3 For more information on the Best of BC Digital Reference Collection, please visit <http://www.gibsonlibraryconnections.ca/bestofbcbooks.htm>.

4 Since public libraries will be able to have access to the collection, everybody can access the titles for research purposes.

the distribution and sales of titles: from the use of digital content to digital technology used to assist in distribution and sales.

Each part of the publishing process is influenced by digitization and will be addressed in the four sections of this guide. The first section “Context” covers the current landscape of digitization and of the trade publishing industry. It looks at the current situation in Canadian publishing in terms of technological changes.⁵ It briefly explores some of the possibilities of digitization as well as some of the constraints faced by publishers. And finally, it explores some of the issues raised by free versus paid content and by copyright.

The subsequent three sections cover the process of establishing a digital strategy, which can be broken down into three major steps. The first step is to establish the motivations for digitization. The section “Why go digital?” addresses some of the key advantages and opportunities of digitization. It looks at website strategies, viral marketing, the use of online browsing and search engines, and e-books. This first step towards digitization is an important one as it demands from publishers to establish the reasons and their needs for digital content, which will influence the digital strategy and the type of digital content necessary for it.

The section “What do you need?” first describes the different types of digital rights needed in author contracts (primary, subsidiary, and marketing rights). The author contracts need to address digitization for publishers to have a sound digital strategy. This section addresses the way publishers can get the necessary rights for their digital strategy. Also covered in this section is the decision-making process for choosing titles to be digitized. Pilot project initiatives, frontlist versus backlist scenarios, content of the titles, and author platforms are all discussed, and questions and considerations for publishers to use as guidelines to select the titles to be digitized are provided.

Finally, the section “What does digitization look like?” discusses the different platforms, formats and digitization processes available to publishers. It addresses issues of control, storage,

5 In the last twenty years, major technological developments have affected the publishing industry. Equally important, digital technologies are continuously improving and being developed.

and management of titles once they have been digitized. It looks at digital asset management and digital asset distribution. It explores the services provided by third parties and the in-house management and storage of digital assets. Also, this section addresses the current lack of standardization amongst formats and platforms, highlighting the importance for publishers to take part in creating standards and to cooperate with organizations that strive for digital publishing standards.

INTRODUCTION

This section on the contexts of digitization and publishing begins with an account of the changes that evolving technology has had on the publishing industry. Digital technology continuously affects the publishing industry and presents many new possibilities for publishers. Yet, publishers also face many constraints and obstacles when they embrace these digital technologies. These possibilities and constraints are explored in this section, with an eye on the fact that the digital and technological future is unknown. Finally, this section addresses one of the greater fears that publishers have when they examine the possibility of digitization: the unattributed and unpaid distribution of content and the need to make material available for free.

CHANGING TIMES

Canadian trade publishers already face many challenges. Why should publishers add another daily challenge and be concerned with changing digital technologies? The publishing industry has been successful for hundreds of years as an industry providing content in physical form through the printed page. Why change a model that has proven to work?

The answer to those questions is that times are changing. The current publishing process is still functioning and adequate: the printing press, the physical warehouses where books are held, and the distribution of books to booksellers and end-readers are all working as they should. However, this physical publishing process is now only one option among several other methods of publishing, distributing, and consuming content. In recent years, the web and digital technologies have impacted people's content consumption and entertainment: we increasingly rely on the Internet to quickly access and provide entertainment. As the technology evolves, our everyday consumption and use of technology also evolves: technology influences how we access and read information. More and more people use different technology, such as the web, phones and PDAs, to search,

distribute, access, and consume information. Newspapers and magazines are a good example of an industry that lived the shift from printed to digital content. Periodical publishers are continuously expanding their presence online, hoping to attract more readers and advertisers, because their print circulation is struggling (Associated Press 2007). Some periodical publishers simply replicate the printed content in digital form, while others create completely new content for the digital version of their magazines/newspapers (publishers will be familiar with this model by taking *Quill & Quire* as an example of the online-and-printed model). Although books are used and read differently than periodicals, the print-to-digital transition by newspapers and magazines does provide a glimpse at content consumed in digital rather than printed form, and some of the tactics used by periodicals publishers to address this issue. The transition from print to digital media is a warning signal for book publishers who have yet to see the shift of technology and consumption of information in their own publishing practice.

Yet skepticism might be understandable as a response to these changes. More than a few trials and setbacks have affected the publishing industry.⁶ Skeptics need to consider, however, that in spite of the past lack of success of technology, book publishing has already benefited from technological development. Darren Wershler-Henry, author and past senior editor of Coach House Books, reminded listeners during his BookNet Canada Speaker Series presentation, that in terms of publishing, the Internet has been very good for “grunt work.” For example, technological tools such as databases used for data exchange and analysis have reduced unnecessary steps in the exchange of information, and have combined data in meaningful ways, providing a greater understanding of the industry to those within the industry (Wershler-Henry 2007). Another example of technology already benefiting the publishing industry is new ways of selling and marketing books through digital technology: online retailers like Amazon have had a major impact on the publishing industry. In other words, technological changes have already affected the publishing industry,

6 The first wave of e-books (approximately ten years ago) did not prove to be as successful as expected. By 2001, Krystina Ross, McClelland & Stewart's general manager, said that "sales ha[d] been less than staggering, ranging between one and 30 copies a month" (Crawley 2001).

probably for the best. Publishers should not fear further digital advancements for the industry, but rather should embrace them; it is absolutely necessary for them to use this new technology. As technology evolves, the publishing industry is invited to follow and adapt to the changing environment, an environment that includes new digital possibilities. This does not mean every opportunity will be advantageous, but rather suggests that publishers need to be aware of changes and of how these changes influence current best practices of publishing. In other words, publishers cannot ignore changes.

Resisting change may have devastating results. As publishers tackle the digital world, recent changes in the music industry can be consulted for guidance to identify errors to avoid and, in cases where things go awry, steps needed to remedy the situation. In the case of the music industry, the problem is the free distribution of content where users download music illegally. For that industry, a credible third party started the road to recovery: Apple, through iTunes, provided users an easy and inexpensive method to download the content desired, within the law. For publishers to succeed where the music industry failed, the publishing industry must be involved in setting processes in place, rather than waiting for a third party to correct problems later on; this will ensure that publishers avoid a loss of control of their content and avoid a detrimental loss in terms of sales. The lesson from the music industry is that publishers should not avoid change, but welcome and control it.

POSSIBILITIES

Digital developments do not need to be risky endeavours where publishers stand the chance of losing everything. Many possibilities can and have emerged from new technologies. The present section examines briefly the possibilities (other than the digitization of title, which will be addressed in the section “Why go digital?”), that have emerged from digital technology.

Evolving digital technology can provide **financial opportunities**, both in term of cost relief and additional revenue for publishers who are pursuing traditional goals, that is to say, creating

print-on-paper books and selling them through bookstores and other retail venues. Digital technologies, such as on-demand and digital printing, can relieve pressures by reducing printing costs for small print runs. These technologies also provide the means to create digital versions of titles; these digital versions become another method to sell content to audiences. Digital versions can increase the publishers' revenues without requiring the high initial costs linked with traditional print production. As such, digital technologies can help reduce costs and increase revenue.

Digital technology simplifies the **distribution** of books, or rather presents an alternative to traditional distribution of books, through the use of digital files. The distribution of content has been accelerated by digital means, whether one considers the distribution within the publishing industry (from publishers to printers, online retailers, and e-book distributors) or the distribution to end-users and consumers. In the case of e-books, digital distribution is not only a rapid method of distributing content, it is also an effective means to cut down on traditional costs associated with distribution, specifically with transportation and warehousing: no large facilities are required. In other words, digital technology helps both increase the speed and decrease the costs of distribution.

Digital technology makes it easier to gather information about book readers and about ways content is used by the readers.⁷ Digital technology, specifically the web, can help publishers identify possible and different **niche markets** and reach them cheaply. As more information is gathered about readers, publishers can focus their efforts and develop appropriate marketing strategies to reach their target audience. The Internet provides access to affinity groups (user-made online communities), and serves as an ideal digital tool to reach and learn about these communities and possible new readers. Book publishers need to understand how online communities function to find the best method to reach them. With digital technology, publishers have the possibility to understand better their current readers, define their target audience, and reach new markets.⁸

7 For example, publishers can use Google Analytics to monitor how people use their sites.

8 For more information on and example of how to reach online niches, refer to "Why go digital?"

Finally, for environmentally conscientious publishers, digital technology can prove to be a friendly alternative to the use of paper. Streamlining the publishing process or pursuing digital formats for books to avoid excess use of paper can alleviate publishers' contribution to the environmental strain. In a more and more "green" world, digital files and digital exchanges are methods to respond to social attitudes; digital technology represents the possibility for publishers to be more **ecologically responsible**.

New financial advantages, distribution ease, the availability of new niche markets, and ecological improvements all represent general possibilities emerging from the availability of developing technology; although they do not exclusively stem from the digitization of titles (i.e. titles being available in digital formats),⁹ they should not be forgotten by publishers as they build their digital strategy.

CONSTRAINTS

Along with opportunities come constraints. Two of the greatest constraints are **time** and **money**. To implement a digital strategy, publishers need to devote time. They can choose to have an in-house employee to coordinate and manage the strategy or they can supervise a third-party service provider to manage their digital process. Finding time is always difficult in an industry where employees are constantly trying to keep up with demands. Making time for "possibilities" or for digital endeavours that might fail does not always seem like the best financial choice. Delegating work to an employee (even an intern) or hiring a third party will also increase costs for publishers. Investing money in possibilities, and not certainties, is a difficult choice for publishers with a limited budget. However, this constraint dissipates when digitization is fully integrated in the publishing workflow; once the digital strategy is implemented and incorporated into the workflow, the process will be efficient and will no longer need the assistance of an employee for the implementation.

⁹ Opportunities specifically from digitization will be addressed in "Why go digital?"

The lack of standardized platforms and formats creates **confusion**, acting as an obstacle for publishers that might be interested in digitization. Too many alternative formats and platforms are used and continue to be developed. Currently, there are different ways for readers to access digital content: from the Internet to digital devices such as e-readers and PDAs. As well there are various digital formats, including images (for example, GIF, JPEG, and PDF files) and flowable content (for example, XML). This overwhelming choice of platforms and formats on both ends of the publishing process means that the communication between publishers and consumers is peppered with additional challenges. Manoeuvring through decisions regarding platforms and formats becomes a new constraint in the process of building a digital strategy. The variety of platforms and formats makes it harder to distribute content to readers.¹⁰

Finally, one of the most important constraints for publishers is the **fear** of being subjected to the unattributed and illegal use of their content, that is, the fear of profit loss from the uncontrolled, free distribution of content, the fear of losing control over digital content, and the fear of not being attributed proper credit for digital content, especially if the digital content is seen favourably within the online sphere. These fears can prevent publishers from truly engaging in opportunities emerging from digital technology. However, all these constraints can be overcome. The benefits of digital technology promise to outweigh the obstacles and constraints.

THE UNKNOWN

Part of the challenge of developing a digital strategy comes from dealing with the unknown. No one can predict today what will happen in the next six months regarding digitization and publishing. What formats will be more common? Which service providers will become necessary? Will readers accept new e-book interfaces and features? Yet, as mentioned earlier, digital technologies are rapidly evolving and book readers are increasingly becoming users of technology to research and consume content.

10 For more information on formats, please refer to the section "What does it look like?"

To survive change one needs to be *flexible*. Getting ready for change requires integrating the capacity to be flexible and to smoothly integrate future change, while still paying attention to current needs. Focusing on speculations about the future of a given book will not be as beneficial as establishing the benefits digitization can accomplish now. In the event that the digital world does not develop as speculated, one's digital practice will not have been fruitless: it will have addressed digital issues of the time.

The good news for publishers skeptical of incorporating digital technology into their publishing process is that they are already using digital technology and are therefore on their way to developing a digital strategy. As Alison Knight points out, "all books begin as e-books" (Knight 2007, 5). Many publishers have knowledge of digital technology such as PDFs, HTML, websites, and metadata, and make daily use of technology. Most publishers use computers and electronic devices on a regular basis to exchange information and data necessary for their publishing practice (for example, contracts, marketing material, and sales information). The process of going digital is not as difficult as it might first seem. Although a digital strategy needs to be open to an unknown future, the steps to develop it are based on familiar grounds with potential for flexibility.

FREE VERSUS PAID CONTENT

One of the most important questions for publishers digitizing titles addresses ways of managing the access to digital content, in order to prevent illegal use of it, and to ensure that the publisher receives payment for all copies used and distributed (within the provisions of the Copyright Act that allow restricted use without compensation). The web has led readers to expect free access to content, an expectation that seems to conflict with publishers' and authors' goals. The answer for publishers has been the promise of extensive DRM (Digital Rights Management) protective measures, but publishers must find the right balance between control and accessibility to avoid alienating both authors and readers, and to maximize their gain (Slowinski 2003, 22).

DRM measures refer to the information technology that is used to control, protect, and manage electronically the work when the content is accessed digitally. This technology facilitates distribution, promotion, and secure sale of electronic content ensuring that the content will not be copied, printed, manipulated, and/or accessed illegally (Slowinski 2005, 23; Knight 2007, 15). Unfortunately, there are some limitations to DRM technology. The more restrictive the DRM measures chosen, the more difficult the files are to process for the e-reader, that is the time to process the information from the DRM takes longer which affects the time to display properly the content for the user (Perry 2005, 254). In addition, DRM measures limit a reader's use of the file: many consumers have problems with the activation of their books if they are trying to use the file on different platforms. In short, DRMs seem to make the simple process of reading a much more complicated process: e-books are harder to read when the reader needs to go through a process to access the content (because of DRM technology); unlike a physical book that simply needs to be opened (Beschizza 2007). This is frustrating for users.

Supporters of free content often talk about users' rights to access and consume content without the constraint of DRM. The notion of user rights is a novel one; one that developed in conjunction with digital books and DRM measures. Some hold that publishers should not be concerned with "user-rights," but publishers should at least be aware of these discussions when making decision regarding the levels of DRM measures they select for their digital strategies. After all, they are an established part of the legal framework defined by the Copyright Act.

DRMs increasingly have a bad reputation with users: DRMs are seen as constraining, imposing, obtrusive, and invasive (when user information is collected), and as preventing the uses traditional to printed books.¹¹ Therefore, when DRM technology is used, publishers should choose carefully to what extent the DRM measures are used (i.e. extremely confining or permit-

11 For more information on DRM and user rights, please read Cory Doctorow's manifesto, a presentation first given to Microsoft's Research Group in June 2004 (available as a PDF on <http://www.changethis.com/4.DRM>). Cory Doctorow is a writer, journalist and blogger, well-known for his discussions on technology and copyright.

ting some flexibility). DRMs should not prevent users from the benefit of the format. In 2003, the Association of American Publishers (AAP) published the results from the study by the AAP's "Enabling Technologies Committee" regarding users of e-books and their responses to DRMs.¹² They found that, at a minimum, the text should remain searchable, the files should have the possibility to be (at least partially) copied and printed for personal use, the text should be accessible to blind and disabled persons, and finally the files should be interoperational (both in terms of format and device) (Slowinski 2003, 58). In other words, readers can accept DRMs, as long as important features of usability¹³ are maintained.

DRM measures can be used to protect publishers' content. However, publishers should also consider providing free content without DRM: some free digital content is necessary to effectively promote and market books. Although most publishers simply want to protect their content from being distributed and consumed for free, publishers should at least consider the possibility of making digital content available without cost. Another important decision for publishers regarding digitization has to do with the choice of the content to be sold and that to be given away (Wershler-Henry 2007).

Free content does not necessarily mean the loss of sales for publishers. There are books available online for free that also sell as hard copies. Some publishers even argue that they sell more copies by providing the content free online than if the book had been only made available for purchase (either as a digital or physical versions) (Kaplan-Myrth 2007). For example, the publisher Irwin Law Inc., publisher of the book *In the Public Interest*, by Michael Geist, said they sold more copies than expected even though the entire book is available as a downloadable PDFs on the publisher's website (Kaplan-Myrth 2007; <http://www.irwinlaw.com>).

12 The published document is the AAP's DRM White Paper: *What Consumers Want in Digital Rights Management*. See Appendix A for the most frequently cited e-book features by users as listed in *What Consumers Want in Digital Rights Management*.

13 The usability of the document refers to the use of features such as printing, copying, and reading the content by users without being too difficult for an average user.

Michael Geist's book *In the Public Interest* (<http://www.irwinlaw.com/books.aspx?bookid=120>) and fiction titles by Cory Doctorow (<http://craphound.com/est/download.php>) are examples of books made available as free downloadable files. Free downloadable books, as a publishing model, seem to challenge the contemporary belief in book publishing that to make a profit, books should be sold, not given (specifically because giving one copy online may result in "too many" copies being distributed). However, free digital versions can be viewed as an extended form of Advanced Reading Copies (ARC). Compared to ARCs, free digital books are not simply review copies, but also can be a way to create something substantial (i.e. an entire book, not simply a blog entry) that is quickly available to a wide audience (since it does not depend on the traditional printing and distribution process).

Free content can be a highly valuable asset to publishers and does not necessarily mean the loss of revenue. First, free digital content is easily shared through viral marketing¹⁴ by being easily shared and distributed across the web. Ideally, free content reaches wide and diverse audiences. Publishers are exchanging free content for online "word-of-mouth" recommendations (Sedgwick 2007). A second method to profit from free content is to attach ads to it. Revenue is generated not from the users of the content, but from the advertisers that benefit from reaching their target audience (O'Donnell 2006b). Ads are still a controversial practice, since they seem to breach the traditional format of the book. But, people are familiar with Google ads on websites and ads have been used in printed books in the past. Advertising can range from Google ads, to ads placed in printer-friendly formatted copies to be downloaded from websites, to short promotional pieces

14 For more information on viral marketing, see section "Why go digital?"

placed before and/or after the first or last chapters or before video and audio digital content.¹⁵ Of course, this is not a common practice within the publishing industry, but it has been in the newspaper publishing industry: content is available free of charge to print, copy, and email, but has contextual advertisement included. The jump for periodicals to use ads with digital content has been more fluid since they already rely heavily on ads for their print medium. For example, the *Toronto Star* online version (thestar.com) provides free access to its content, however, in the case where a reader wants to print the article, the printer-friendly format includes advertisement:

[L]inks [...] appear at the top and bottom of each article and they [...] ride along with the article no matter where users send it. The links support each other by providing integrated cross-sell and upsell between free uses, reprints, and fee-based licenses (icopyright.com).

For publishers who want to remain ad-free, a third option to distribute free content comes from collecting user information. Free content can be used as an incentive for users to sign up for newsletters and member-only sections of a website. This information on users (name and email addresses, and if using a login system, page views and click-throughs) is valuable as it gives publishers direct contact and knowledge about their audiences. A member-only section on the publisher's website means that the publisher can monitor who accesses which excerpts, at what

15 Although no trade publisher is currently using ads with their digital content, I do not consider such an act to be speculations about the future of digitization because the technology necessary to use advertisement is available today. To be clear, the use of ads is not a suggestion to transform books into magazines, but rather to look at limited and targeted use of advertisement in exchange for free digital content. In addition, it is important to remember that giving free access to digital content is a form of promoting (i.e. advertising) the book. Ads are not necessarily meant to be part of a digital book: ads can accompany other forms of digital content related to the book. For example, the digital content can be a video clip or a podcast.

To date, the closest form of ads and book content is with Google Book Search, where once the content is viewed on the user's screens, Google AdSense are listed at the bottom of the screen.

frequency, and determine which subjects are more popular. For example, Random House Canada created BookLounge.ca where readers can become members of the site, giving them access to more digital content. The membership is free for users, but it provides the publisher with a database of their customers and information about their readers.

Free content is especially valuable to reach potential readers, yet publishers often fear that free content will not be credited to the publishers and authors. But free access to part or all of the content of a book, to a selection of photographs, to a podcast, or to a video interview does not mean that no copyright is enforced for the material. This is precisely why initiatives have been created to give the opportunity to distribute free content and reach greater numbers of customers while still maintaining copyright. Such an alternative is the Creative Commons licenses (www.creativecommons.ca). The use of Creative Commons licenses ensures that readers are aware of the copyright terms for the free digital content, and aware that certain conditions for the use and distribution of digital content apply even if the content is free (www.creativecommons.ca). Creative Commons licenses make copyright visible to users, not just as something imposed unto them. However, Creative Commons licenses are not DRM measures: they do not use technology to prevent users from copying, printing, saving, and sampling content. Publishers must trust that users will respect the Creative Commons licenses. At the same time, the purpose of free content with Creative Commons licenses is to encourage the distribution of the content: the goal is to facilitate that content to be passed along, copied, printed and sampled, in order to create online word-of-mouth. Creative Commons licenses set the guidelines for this free distribution. To adapt to the different needs of the distributed content, Creative Commons has four different licenses that can be mixed and matched. These deal with the attribution of the work to the author, the commercial use of the work, and the possibility to use the work in other works.¹⁶

Creative Commons and similar initiatives can be used by publishers not only on entire books, but also for marketing materials or for selections of the work (assuming that the appropriate

16 Refer to Appendix B for the breakdown of the four Creative Commons licenses.

rights are licensed under author contracts). Their goal is to ensure that copyright information travels with the free content and that users have a better understanding of copyright. By using Creative Commons, copyright information is maintained and users are encouraged to share the content (under the restrictions) thus promoting the work. With initiatives such as Creative Commons, the content circulates more freely amongst the audience, an extremely valuable element for marketing purposes. Free content is useful for the promotion of books, which potentially results in increased hard copy sales. With Creative Commons, publishers can benefit from free content as a marketing tool while they can still maintain some control over the content and be reassured that it will be attributed to the author and publisher.

SECTION CONCLUSION

To create a sound digital strategy, understanding the context of the publishing industry and digital technology is important. As we have seen, publishers approaching digitization must engage with developing technology; publishers must adapt. However, adaptation does not serve only as a means of survival; these digital changes can provide many possibilities for publishers. Change comes with certain obstacles and constraints, namely the amount of time and money needed, the confusion regarding platforms and formats, and the fear of losing profit from the illegal distribution of digital content.

The issue of free content brings up questions regarding the protection of content, both in terms of actual files and in terms of copyright. Although publishers may not consider user expectations as a primary concern, they should pay attention to some of the possibilities free content provides. Free content can be a promotional tool to sell more hard copies by creating word-of-mouth and can supply other possible revenue through advertising.

WHY GO DIGITAL? (ADVANTAGES AND MOTIVATIONS)

INTRODUCTION

As briefly discussed in the introduction, many opportunities are available for publishers who develop a digital strategy for their titles. This section focuses specifically, given the broad realm of digital technology, on the advantages and motivations for publishers to digitize titles. The opportunities and advantages discussed here center on current trends, and do not represent speculations about the *future* of digitization. At present, the greatest advantage of digitization is its impact on the overall marketing and promotion of books. Websites, viral marketing, book browsing and searching, and e-books (in this context, as digital versions of the title) represent key advantages and therefore motivations for digitization. This section first looks at website strategies (for publishers, authors/books, and bloggers/social networks) and at ways these can be used for marketing purposes. Second, it addresses viral marketing and some of the tools publishers can use to their advantages for promotion. Third, this section looks at book browsing and searching as an important method of familiarizing readers with the publisher's titles and of marketing books online. Finally, we examine e-books, both as digital versions of a title and as marketing devices themselves.

WEBSITES STRATEGIES

The digitization of titles is important since it produces the content that will populate websites to help promote book titles and reach readers. Websites are one of the best methods to attract attention to a publisher's list. Shannon Maughan reports in *Publishers Weekly* that:

most publishers concur that the broad exposure online marketing is bringing to books and authors is invaluable, and that more often than not, there is a strong

correlation between the sales of books and the online exposure those titles receive (Maughan 2007).

For publishers to benefit from online marketing efforts, they need to take an active part in the online world to promote their titles. The following addresses the need for publishers to use the different websites available to market their books and reach readers to ultimately increase sales. This section looks at three types of websites: websites of publishers, websites created for specific authors and/or books, and blogs and social network websites.

Publishers' websites

As previously mentioned, to help promote books, publishers must use online marketing tools; one of those tools is the creation of websites for the publisher. Most publishers already have websites, but these websites are often not created for the reader. Publishers' websites tend to target book retailers and libraries: this is a limited use of publishers' websites' potential.

To promote books and reach readers successfully, publishers' websites must provide up-to-date information about their authors, books, and events. Too many publishers' websites only provide basic information about the titles (such as ISBN number, publication date, price, and BISAC categories) instead of appealing to the readers' desire to find out more about books and about authors (such as content and stories, and not metadata). Information needs to be relevant and interesting for readers: digital content on websites should derive *from* the titles and authors (for example, excerpts from the books, photographs and illustrations, and author interviews).

Publishers must also engage and interact with their (potential) readers. Rather than simply feeding information to readers, publishers must start interacting with their audience. For example, for fall 2008 Penguin plans to create a song-writing contest where readers can submit their songs and then vote for the best one. The song-writing contest will be tied-in with the book *Audrey*,

Wait! by Robin Benway (Sellers 2007).¹⁷ This contest engages the readers with the publisher and the book, creating “buzz” for a title. Sponsoring contests is only one of the ways to interact and engage with readers. Following are other methods that publishers can use for their own websites:¹⁸

- Provide an online space (a forum) where users can have the opportunity to leave comments or notes, write reviews¹⁹ of books they read, and interact with other fans of the publishing house/authors/books.
- Update the content on the website with new information surrounding the books and authors (such as new awards, new titles and authors, new tours, author “gossip,” information on books that relate to contemporary topics, or topics in the news)²⁰
- Create multimedia content (such as videos, podcasts, and slideshows)
- Make each page of your site “linkable” (i.e. each page has a distinct address that can be easily shared between people)
- Make some “behind-the-scenes” information available (this is the online and publishing version of DVDs “special features” like bloopers, deleted scenes and actor interviews).
- Link to other websites with information you believe is interesting and relevant to your own material.
- Provide free downloadable content (in limited amount), such as excerpts, photographs, videos, and podcasts.

Publishers’ websites depend on digital content: this content is made possible by digitization. Without digital content, publishers lose the opportunities to successfully use websites as marketing tools to reach readers.

17 The song-writing contest ties-in with the book because in the book, after dumping her boyfriend, the protagonist’s ex-partner writes a song about the breakup.

18 This is not a complete list, but a departure point.

19 Henry Holt Books, Simon & Schuster, and Random House work with their young readers in exchange for ARCs: teenagers review books online throughout the year in exchange for books (Sellers, 2007).

20 This is not just a time and place to boast about the publishing house, but this is any tidbit of information that makes the readers say “cool!”

Websites created for authors and books

This section explores websites that are created specifically for individual books and authors.

Although it is important for publishers to have a strong online presence of their own, readers most often recognize authors and book titles before they recognize their publishers. Publishers have started to build small and targeted sites for individual authors and books. For example Douglas & McIntyre has created a site for the book *Trading in Memories* by Barbara Hodgson (<http://tradinginmemories.com>). This website has detailed information about the author and the book, an interview with the author, excerpts, and reviews. The content of the website can help create the “back story” of the book and author; it helps to solidify the relationship between readers and authors. As author Barry Lyga (*The Astonishing Adventures of Fanboy and Goth Girl*) states about his website: “[the website] helps to cement the relationship you have with the reader” (Maughan 2007). The strong bond between reader and author translates into loyal readers, and hopefully loyal customers.

Websites dedicated to authors and books must contain substantial amount of content and must be updated on a regular basis to successfully reach and interact with readers. A site that is updated will invite readers to continuously return to the site. Creating a site that becomes a destination for readers strengthens the relationship between reader and author (and subsequently, with the publisher as well). This bond is useful to promote books successfully. Up-to-date websites and online marketing help not only the promotion of books, but they also serve the purpose of keeping and expanding the established readership around the authors, even in between books (Johnson 2007b).

Blogs and social networks

Sites by third parties can also be used to promote titles and authors. Given that third parties appear to be more objective and less commercial, they have more impact when recommending titles (Seller 2007). Many bloggers review books: while there are bloggers that write specifically

about literature and books,²¹ others focus on particular subject matter, such as cooking, cities, technology, gardening, history, arts, etc. Both types of bloggers should be contacted by the marketing department when titles are ready to be promoted. Marketing staff can send bloggers ARCs, catalogues, press releases, and any digital content available, like links to multimedia content (audio, video, photographs, etc.). To select online book reviewers and bloggers, publishers should ensure that the bloggers/reviewers meet the needs of the marketing department to reach a targeted audience. Consider the following questions:

- Does the blogger review books similar to those on the publisher's list?
- Have they reviewed a title from the publisher without being asked to do so?
- What type of (digital) content do they use and request for their sites ? Do they welcome/request ARCs and/or book catalogues?
- Does the blog receive paid advertisement? How much do they charge per month? This can help indicate the popularity of the blog with viewers. Some sites will provide a media kit, which often lists statistics regarding their average page views and unique visitors per month (i.e. how many people look at the site, and advertisement, and how often)
- Which blogs do the authors read? (Authors can help determine which blogs are appropriate for their books: the authors might already have an established relationship with bloggers, and they might comment and interact on the blogs) (Mandel 2005).

Choosing bloggers is similar to selecting appropriate print media to promote a title (i.e. selecting magazines based on their subject matter, sending ARCs to book reviewers, and contacting publicists). To find these bloggers, publishers must invest time to search online, but publishers can also get references from their authors as a good departure point.

21 For examples include Book Slut (<http://www.bookslut.com>) and Book Ninja (<http://www.bookninja.com>)

Other than bloggers' sites, publishers can use established social network and community sites to reach people. Sites such as Facebook,²² MySpace, Flickr, and YouTube are all sites where people create their own communities and freely exchange information and recommendations. Access to these sites provides an entry to large groups of readers. To have a presence on these social networks, publishers must create their own profiles (i.e. a page describing themselves) on these sites, where they can promote and recommend (i.e. advertise) new titles and events (book launches, author tours, and festivals). Social networks are different than the traditional publishers' websites because first, they do not hold the same detailed description of the content and second, other online users of these sites can become "friends" of the publishers or "members" of the publishing "group." As friends and members of the publishers, users not affiliated to the press, but friends of the publisher's friends can access the publisher's profile (hence the reason for the name "social network"). The online profile of the press can be viewed by people not even searching for the publishers books and content. Publishers like Arsenal Pulp Press (<http://www.facebook.com/group.php?gid=2353427989>), Coach House Press (<http://www.facebook.com/group.php?gid=2260058751>), and House of Anansi Press (<http://www.facebook.com/group.php?gid=2260468419>) have all created accounts and groups on Facebook to reach readers and promote books and events. Even large publishers, such as HarperCollins Canada, use Facebook to promote their books. Last May, HarperCollins Canada decided that, instead of simply advertising the author tour of *PostSecret: Extraordinary Confessions from Ordinary Lives* by blogger Frank Warren in traditional media, they were simply going to create an event page on Facebook notifying users of the event happening in Toronto. Approximately 1,800 individuals joined the group and, as quoted in *Quill & Quire*:

22 Facebook is recently getting a lot of attention as an important place for online marketing. For more information on the different tactics and opportunities available on Facebook, please visit "The Facebook Marketing Bible: 24 Ways to Market Your Brand, Company, Product, or Service Inside Facebook" on *Inside Facebook* (<http://www.insidefacebook.com/2007/12/09/inside-facebook-marketing-bible-24-ways-to-market-your-brand-company-product-or-service-in-facebook>)

“People went mad,” says HarperCollins publicist Barbara Bower. “It ended up being a wild, wild success” (Johnson 2007a)

All these different websites (those of publishers, authors, bloggers, and social networks) have the same basic need. They all need digital content that appeals to readers. The digital content for online marketing purposes is not limited to a digital version of a title, but also includes enhanced digital or multimedia digital content. The content of websites must engage with readers, both in its form and in its material. Here are some examples of digital formats and the type of material that can go online:

Table 1: Examples of digital formats and the type of material that can go online

Example of Digital Formats	Example of Material as Digital Content
Audio (podcast)	Author “backstory” (Interviews with authors, news clips from TV programs, awards, etc.)
Video (short clips; DVD excepts, news reel)	Book “backstory” (Information about the creation of the book; if fiction, more information on some characters)
Images	Book artwork (photographs/illustrations)
Text	Information and photographs about events
	Excerpts (text or audio)
	Contests

On the Douglas & McIntyre website, the non-fiction title *The Last Wild Wolves* by Ian McAllister²³ has a slideshow of a selection of the stunning photographs from the book, as well as a video clip of the last wild wolves (<http://www.douglas-mcintyre.com/book/9781553652427>).

23 *The Last Wild Wolves* is a coffee table book with many large-scale photographs.

However, for a fiction title like *The Silent Raga* by Ameen Merchant, which does not have any artwork, Douglas & McIntyre decided to feature its author and writing: readers can read on the website an exclusive interview with Merchant as well as an excerpt (<http://www.douglas-mcintyre.com/book/9781553653097>). Similarly, for the book by Paul Myers, *It Ain't Easy: Long John Baldry and the Birth of the British Blues*, Douglas & McIntyre fittingly decided to focus on an audio format and asked Myers to create a podcast introducing the book (<http://www.douglas-mcintyre.com/podcast/4>).

In the three Douglas & McIntyre examples, the digital content is crucial to help market the titles. This digital content was a combination of existing content from the books (photographs and excerpts) and new content based on the books and authors (video clip, podcast, and author interview). The material is interesting and engaging, which adds value to the website as an interesting source of information for books. Websites should not be simple replicas of print ads: they need to provide new engaging information. Digital content is key for online marketing and reaching readers (Macfarlane 2005).²⁴

The use of enhanced digital content is not limited to the web. Publishers have started to include multimedia elements within books to market them. For example, HarperCollins Canada offered a limited-edition CD of Indian music with the novel *Sacred Games* by Vikram Chandra to promote the book (Williams 2007). Large publishers like HarperCollins are not the only ones doing so; musician/author Kid Koala worked with ECW Press to create a soundtrack for his graphic novel *Nufonia Must Fall*, and Anvil Press produced a CD to go along *The Beautiful Dead End* by Clint Hutzulak (Williams 2007).

In addition to CDs, HarperCollins started to add videos of authors as part of original e-books (through their imprint Avon, publisher of romance novels and erotica). At different moments throughout the e-book, the author talks about the plot and characters. The original erotica e-

24 It should be noted that publishers must have the necessary digital rights to create the digital content for marketing purposes.

books are then cross-promoted and marketed on erotica sites like nerve.com and Romantic Times (Publishers Weekly 2007a)

VIRAL MARKETING

Along with websites, viral marketing is an important part of an online marketing strategy. This section looks at what constitutes viral marketing, and reviews some of the sites and strategies used by publishers. Viral marketing is defined as marketing online that depends on web users to share and distribute content in order to create brand awareness and/or online word-of-mouth; as the name suggests, the marketing piece or content is distributed “virally,” as in rapidly and through a web of people.

Viral marketing is ideal for publishers: it is inexpensive and engages its audience (it depends on users’ involvement and interaction with the digital content). Social-networking sites work well with viral marketing: sites such as Facebook, where communities are first built on friendships made in the physical world (as opposed to the online virtual world), are places where people share their daily lives, their impressions and reviews, and where they can also share digital content (by linking to other sites or embedding content) (Johnson 2007a). Social-network sites are an example of what is now known as “Web 2.0,” where the focus is on user-interactivity rather than on consumption of static content. These online interactive users are the ones primarily responsible for information-sharing and for online recommendations (i.e. viral marketing).²⁵

The concept of online communities is not exclusive to social network sites; online retailers also create online communities such as reader clubs, where profiles and recommendations are exchanged. Companies, such as Amazon and Chapters/Indigo, encourage readers to write reviews and comments, to list books they recommend, to add titles on their reader-personalized bookshelves, to add “tags” to help other readers find books, and to create networks within the retailer’s

25 For more information on Web 2.0 and advice on benefiting from it, please refer to O'Reilly's book *Why Web 2.0 Matters and How You Can Make the Most of It* by John Musser and Tim O'Reilly (2006). Download an excerpt at <http://www.oreilly.com/radar/web2report.csp>.

site where authors, artists, and readers can connect (amazon.com; www.chapters.indigo.ca).²⁶ Reaching these online communities is crucial for publishers because they are important voices in the realm of online word-of-mouth. Publishers must build a strong presence online to interact with these communities.

An increasingly popular method of sharing digital content is through “widgets.” Widgets are packets of code that can be embedded into HTML-based websites that can add an interactive component to the page (Wikipedia 2007). In some ways, they work as little websites embedded into a bigger one. These are ideal for inserting small interactive ads on websites; however, most importantly, they can be used to safely share book information, without risking that the content will be downloaded, printed, or copied. For example, they allow users to sample book pages and artwork. Since widgets are packets of code, the codes can be easily shared and placed on various websites. Anyone interested, such as book-and-literature bloggers, booksellers, and libraries, can embed the code unto their own websites, promoting the publishers’ books.

Both HarperCollins and Random House created widgets to virally promote their titles. While HarperCollins’ “Browse Inside” widget was made available just days before Random Houses’ “Browse and Search” widget, HarperCollins only offers approximately 2,000 titles compared to the 5,000 titles of Random House. In addition, HarperCollins’ widget only offers sample pages, while Random House managed to provide searchable content²⁷ and audio²⁸ as additional features. Access to sample pages and searchable content increases the likelihood readers will purchase a book.

26 Indigo has even created its own version of Facebook, simply called “Community” (http://www.chapters.indigo.ca/Community-Home/community_home_promonb.html).

27 More on the search function of widgets under the sections “Book Browsing and Search Engines”

28 Random House has audio widgets, which function like the text widget. A person can simply cut and paste the code on their own sites to have access to the different audio files. Each audio file is a two-to-ten minute sample of authors reading excerpts of the books (<http://www.randomhouse.com/audio>). The audio widget also has a link for readers to order physical copies of the audio book or download the audio digital file (http://www.randomhouse.biz/webservices/insight/widget/userguide#add_audio)

Widgets are a great example of viral marketing and give publishers another opportunity to reach readers online.

Finally, HarperCollins' widget is clearly branded with their logo, while Random House's opted for a more subdued look and logo-less widget. The lack of branding seems to appeal more to bloggers who prefer to avoid direct affiliation to a corporate identity (Reid 2007b). The decision of branding the widget or not is a decision that must balance the desire for brand recognition on the part of the publishers and the invisibility of the brand that helps dissipate the negative connotations of corporate identities and subjective recommendations. However, the importance of widgets comes from the fact that the contents of books (sample pages, covers, etc.) are not just located on the publisher's website, but throughout the Internet. Widgets increase the visibility of the books online. They are a prime example of viral marketing made possible thanks to the easy exchange of code, and therefore of content.

The following are tips to reach readers through websites and viral marketing:

- “Re-energize” your website by adding new content regularly (Maughan 2007)
- Encourage authors to blog and create their own accounts on social networking sites like Facebook, MySpace, Flickr, and YouTube
- Create a publisher’s profile on social networking sites and promote new books, events, and authors. Ask staff members to collaborate and add content regularly on the publisher’s pages on social networks
- Create sites for authors and books
- Create links between websites and social networks
- Use Web 2.0/interactive content: create and share podcasts, videos, and images; place banner ads on pertinent third-party websites (such as literature bloggers); create book trailers; use contacts with bloggers (sending them ARCs, catalogues, and digital content (such as videos, podcasts, and photographs), upload multimedia digital content on social networking sites; and develop contests for readers online.

- Encourage authors to both develop back stories for their books (extending the story, especially for fiction titles, keeps readers “in the know” and connected to the authors, strengthening the author-reader bond) and share information about themselves and their writing process.
- Keep yourself in the loop and updated on new developments about the digital world: What are people doing online? How are they interacting and communicating? A good place to start is by regularly visiting technorati.com/technology (the technology section of technorati, an aggregator of all the content of blogs) to keep on top of new digital technology.

BOOK BROWSING AND SEARCH ENGINES

Amazon and Google have been an important online presence in publishing. Both companies have created programs to work with publishers to market and make accessible publishers’ backlist and frontlist titles. These two companies have made book browsing online and book searching a common practice.

Book browsing

Selling books online is nothing new. Amazon has been a key player in the sale of books online since its launch in 1995. With “Search Inside the Book,” Amazon gives online purchasers a peek at the book as it is printed.²⁹ This online browsing is believed by some publishers to make the purchase of the book more likely: it gives readers a sense that they *know* what they were purchasing (the look and content of the books) (Skelton 2005). Online browsing gives the impression of looking through the book by displaying JPEGs of the cover, of the table of contents, and of the index. Google Book and book retailers with online stores, such as Barnes & Noble,³⁰ now have programs similar to Amazon’s: they too provide the possibility to look at books online.

29 Amazon’s “Search Inside the Book” feature started in 2003 with 120,000 titles.

30 Barnes & Noble is also jumping in the game of browsing. On 16 October 2007, they announced the launch of the digital book preview “See Inside,” a feature created by Zinio (Maul 2007).

Online browsing seems to be one of the most important tools for online retailers to sell hard copies of books because of the preview it gives to readers. Making the files available on one's own website or on online-retailers' websites is necessary. Publishers can place JPEGs of the covers and inside spreads as thumbnails and as larger images on their own websites as a method to preview the books, and files can be sent as PDFs (of the text) and JPEGs (of the covers) to Amazon, Google, and book retailers. At a minimum, a publisher should provide readers with photos of the covers; however, giving readers the possibility to truly browse the book can make a difference in sales (LibreDigital 2007, 6).³¹ People hesitate to buy what they cannot see. Though there is no tactile experience online, no flipping and touching of pages, publishers can provide readers a visual browsing experience:

[I]t has become increasingly important to provide the best buying experience possible online. The logical first step is to offer online browsing inside the book itself, especially trade titles (LibreDigital 2007, 6).

Some publishers, like HarperCollins ("Browse Inside") and Random House ("Browse and Search") have started to incorporate online browsing on their own websites. Rather than depend on a third party whose browsing capacity is limited to one web domain, these two companies created their own widgets to provide their own online browsing experience, that is also a viral-marketing tool. As mentioned earlier, these widgets can be incorporated onto their own websites and on any HTML-based sites. Most of these browsers offer a feature that lets readers have access to more than the covers. Both Random House and HarperCollins believe that widgets will help readers discover books online (Grittani-Livingston 2007a; Deahl 2007).

31 Based on their own statistics and research, Amazon states that their "Search Inside the Book" browsing function increases sales (PW NewsLine 2007).

Search engines

Google, Amazon, and Random House not only lets users browse the books, but also incorporates a search function.³² Search functions work well with online marketing.

Most book search-engines work by having access to the entire content of the book; the engine searches for particular words and highlights them when found inside the text. Only limited, surrounding text (a “snippet”) is shown around the searched words. Once all the instances of the words are listed, the reader can click through the list and view a specific page where the terms are located. There is also the possibility of simply browsing the book by clicking through the pre-selected pages supplied (Amazon.com). In other words, readers never have access to the complete digitized book (even though the publishers submitted the entire book to Amazon’s and Google’s programs, a requirement to take part in the programs).

Google

As many publishers are aware, Google is a major player in book search. The Google Book Search has a Partner Program that gives publishers the opportunity to give their books to be digitized and accessed online.³³ Unlike Amazon, Google is not a book retailer; it simply provides (part of) the book content. To purchase the books previewed on Google Books, users³⁴ can click on a link next to the book featured to have access to the publisher’s website or those of online retailers (Google 2007, 7). Google is trying to build a universal digital library and requests from publishers that they submit the entire book. To work with publishers to obtain all this content, Google lets the publishers decide how much content can be shown to viewers for free (from 20 to 100 percent).

32 Random House’s widgets permits users to search inside the text, while HarperCollins’ widgets only permits browsing.

33 Publishers need to send digital versions of their titles.

34 Users range from researchers looking for precise information to readers searching for a particular book, and to general web users browsing online.

Google's program is free for publishers to join, and any ad revenue generated by Google Ads sales is divided between Google and the publisher. On the web pages where books are displayed, Google places ads (Google Adword) at the bottom of the screen. Revenue is generated when a user clicks on the ad.³⁵ However, though joining and having Google scan the publishers' titles is free, the digital files scanned belong to Google; the publishers do not have access to them. For that reason, it is recommended that publishers keep control of the digital files and provide the files to Google (Davy 2007, 22). Once Google receives the files, the files are indexed and made available on Google Book Search and Google's general search (Google 2007, 7).

Google is useful for publishers not only for its browse and search functions, but also for the information Google gathers about user-traffic for the books on the Google sites. This information provides greater knowledge to publishers about their audiences and the online marketplace (Davy 2007, 21). It gives publishers access to precise reporting on the different publishers' titles; including reports "on the book views, impressions, ad clicks, Buy This Book clicks, and ad revenue" generated by the publishers' accounts (Google.com). In addition, the Google Book program can benefit publishers by driving user-traffic through links from Google Book to publishers' websites. The reports provided from Google Book in combination with Goggle Analytics,³⁶ can help publishers identify their market and their users (Google.com).

Many publishers and authors would probably agree that a fate worse than piracy is obscurity (especially when you are not a mega-publisher). Since "search[ing] is a discovery tool for readers" (Kelly 2006, 9), it is understandable that smaller or more obscure publishers and authors can benefit from Google's programs. To truly exploit these services, publishers should ensure that the files they digitize are flexible; that they have the capacity to include metadata; and that they can

35 For more information of Google ads, please visit Google Adwords (<https://www.google.com/adsense>) and Google AdSense (<https://www.google.com/adsense>)

36 Google Analytics is Google's tracking system. It tracks and analyses how users find and navigate a website and how they become customers (<http://www.google.com/analytics>).

be used on different platforms.³⁷ The digitization of books is crucial for publishers to take part in projects and marketing efforts that use book browsing and book searching as integral parts of the services.

Following are some tips to benefit from online book browsing and book searching:

- Create an account with Google Book and Amazon’s “Search Inside the Book” programs
- Create an account with Google Analytics
- Create digital content for browsing and search: JPEGs of covers and selected spreads; PDFs of inside text for search programs like Google and Amazon
- Avoid simply sending cover images to Google and Amazon; instead, also send entire digital versions of books
- Investigate creating your own widget or licensing it from a different company (for example, Bloomsbury announced last April that it would work with LibreDigital (i.e. HarperCollins) to create their own version of search functions, “Look Inside” (Publishers Weekly 2007b)³⁸
- Include JPEGs of covers and inside spreads on your own websites, not only as thumbnails, but also as larger images.

E - B O O K S

E-books have been on publishers’ minds for at least ten years. Given that the first wave of e-book interest quickly died down, it is understandable that some publishers are skeptical about the second wave of e-readers and the regenerated anticipation. From a publishing industry perspective, e-books have been an effort led by scholarly publishers (Shatzkin 2006, 1).³⁹ For trade publishers, the benefits of e-books do not currently outweigh the obstacles. The lack of standard-

37 For more information on file format, platform and file flexibility, please refer to the section “What does digitization look like?”

38 Right now, there are no small-to-mid-sized trade publishers who have created their own widgets.

39 The low production cost of e-books along with the smaller audiences of scholarly publishing make e-books perfect for scholarly publishers (Shatzkin 2006, 1).

ization in formats and platforms has been a major deterrent for trade publishers to venture into the e-book market.

Yet, e-books are starting to become a growing part of the efforts by large trade publishers to reach new audiences. New e-readers are slowly becoming more appealing and have the advantage of making books easily distributed and accessible. In addition to e-readers, consumers are increasingly starting to read e-books on diverse portable platforms, such as PDAs, computers, and cell-phones, opening access to e-books. *Publishers Weekly* (PW) and *TechCrunch* (TC) have reported the use of cell phones in the UK (PW) and Japan (TC) to read electronic books. In the case of publishers in the UK, due to flood of cell phones and the increasing use of them to surf and purchase online, UK publishers hope this will also mean the increase of e-book consumption on cell phones (Russell-Walling 2007). As Jim Green, digital development director from HarperCollins stated, "the UK is way ahead in mobile phone technology, and in thinking about how it can be used to deliver content" (Russell-Walling 2007). In fact, the hard copy of *Collins Guide to British Birds* will be released with a digital version on a SIM card for cell phones: using one's cell phone as a reference is believed to be more convenient than carrying and using the physical book (Russell-Walling 2007).

E-books currently represent only a small increase in the total number of book sales. Between 2002 and 2007, the United States e-book revenues increased from approximately \$1.5 million USD to over \$8 million USD (IDPF Wholesale E-book Sales Statistics). Though it is still only a small fraction of total book sales and cannot yet be described as a trend, this increase is important as it points to a possible new acceptance of e-books by customers: therefore publishers should at least keep close attention to future developments (developing e-book trends or developing new digital technology⁴⁰).

40 For example, Amazon unveiled last month its own e-reader, the Kindle. Only time will tell if consumers will readily adopt it as a platform to read e-books (Reid 2007a).

Only a few trade publishers have embraced e-books completely. Harlequin and HarperCollins are two examples of trade publishers producing e-books. Harlequin has increased its production of e-book titles from nine titles a month in October 2005 to approximately one hundred titles a month in May 2007 (Vallik 2007, 7). Their e-books range from frontlist titles (approximately sixty titles) to a selection of backlist titles (Vallik 2007, 7). HarperCollins has also started to digitize its collection and offers readers the possibility to purchase and download titles directly from their websites and other retailers' sites. To encourage e-book sales, HarperCollins discounts the e-book by 20 percent of the hardcover price ticket (<http://www.harpercollins.com>).

The performance and trends of e-books can provide clues for smaller trade publishers. The fact that Harlequin is continuously increasing its e-book production and that HarperCollins selects its romance novels for e-books indicates that romance novel readers are amongst the most welcoming of e-book technology. For trade publishers outside the romance market, e-books probably may not increase the publisher's revenue at present, but it still remains important to digitize entire versions of the titles to join marketing programs like Google Book Search and Amazon Search Inside the Books. The digital version of books, available through Google and Amazon, reach different audiences than typical e-book consumers; these complete digital versions of books act as a marketing tool. E-books should not be seen solely as digital versions of printed books read on e-readers, but also as key elements of a marketing strategy.⁴¹ Therefore, it is important for trade publishers to consider digitizing entire books, rather than simply parts of it, in order to have greater flexibility for their entire digital strategy.

The following are tips for e-books and creating digital versions of titles:

- Digitize the entire book, not only a selection.

⁴¹ While many use the term "e-books" to only refer to digital books read on e-readers, for the purpose of this guide, I use a more general definition of e-books to include all digital versions of an entire book title.

- For frontlist titles, integrate the digitization process in the in-house workflow (by discussing, with both the production and design staff, how to integrate and streamline the creation of digital files as the books are produced)
- For backlist titles, choose titles that can benefit from online marketing⁴²
- Submit titles to Amazon and Google
- Submit titles to pilot projects (for example, the Best of BC Digital Reference Collection).

SECTION CONCLUSION

For publishers, creating a digital strategy first requires discerning the different opportunities and advantages available from the digitization of titles. This section looked at different online marketing possibilities as a primary reason for publishers to go digital. First, this section looked at creating and using various websites, populating them with digital content targeted towards end-consumers. It also included information regarding the creation of enhanced digital content such as multimedia files, both audio and video. Second, this section addressed the importance of viral marketing practices, such as social networks and book widgets. Third, this section discussed the importance of book browsing and searching, as important aspects of online marketing and of developing a digital strategy by increasing the exposure of books online. Finally, the last motivation to go digital is the development of e-books. E-book production is increasing, specifically for large publishers, and can represent additional revenue.

42 For more information of book selection, please refer to the section "What do you need?"

W H A T D O Y O U N E E D ?

(D I G I T A L R I G H T S A N D T I T L E S)

I N T R O D U C T I O N

Once the decision to digitize titles and to develop a digital strategy is made, the next step is to go through author contracts to establish the rights the publisher currently holds and the rights necessary for the digital strategy. The choice of titles for digitization can be done simultaneously as the choice will depend in part on the *availability* of the necessary rights. This section first covers the various types of rights needed for a digital strategy: primary, subsidiary, and marketing digital rights. Second, it discusses the issue of competing works within the context of digital rights. Third, it looks at how to obtain the rights necessary for digitization. Finally, it outlines the issues for publishers to consider when deciding which titles to digitize.

T Y P E S O F R I G H T S

Karen Gilmore, Vice-President Operations, Corporate and Legal Affairs at Douglas & McIntyre, reminded publishers during a Literary Press Group (LPG) panel discussion “Rights and Contracts in the Digital Age” (4 May 2007) that to implement and profit from digitization, the publishers must acquire all needed digital rights. Since contracts often pre-date technological and digital advances, some contracts will not mention digital rights or have incomplete and/or unclear definitions of the necessary digital rights. Publishing contracts deal with copyright, which is a package of different rights. Publishers need to clearly identify and define each of the different rights in their author contracts (Gilmore, 2007). In addition, Gilmore stated that the term “digital rights” often applies to two different types of digital content: first, the rights of a *digital version* of a work, and second, the rights for enhanced or multimedia versions.

Gilmore says that the current state in the industry is to consider the digital version of the text as a primary right, and the enhanced/multimedia digital content as a subsidiary right. In addi-

tion, the sale of the digital version of a book is differentiated from the use of a digital excerpt as a marketing tactic (Gilmore 2007). Although the actual files may be similar, the rights licensed under the author contract specify the *purpose* of those digital files. All digital rights are important to obtain for publishers to have greater flexibility with digital content. The following paragraphs look at the three main types of digital rights and how variations can affect the possibilities granted with these rights. *Please note, the following section is not legal advice; publishers must review their entire contracts with their legal counsel.*

Primary rights

The grant-of-rights clause is a *primary* right. This clause should include rights for verbatim digital versions of the title. The following is an example of a clause of rights granted. A breakdown of the information the clause contains is then provided.

The Author hereby grants the Publisher the sole and exclusive right and license, for the legal term of copyright, to produce, publish and distribute, and to license the production, publication and distribution of, the Work or any abridgement of the Work or any substantial part of the Work, in all languages throughout the world, in book and/or print form and by any other means of distribution or transmission, whether now known or hereafter developed, intended to make the Work available in visual form for reading (including without limitation “e-book” format and other electronic or machine-readable media, or on-line or satellite-based data transmission) (Gilmore 2007).

Formats

As part of the rights granted, the publisher should specify that the format includes all book format, that is, printed format as well as any format that is meant to be read by any visual means. This ensures that all print *and* digital versions of the book are included.

“[...] in book and/or print form [...] available in visual form for reading”

Transmission/distribution

The methods used to distribute and transmit the content of the work should also be defined in the rights granted. E-books formats and e-readers, electronic medium, online platforms, and data transmission by satellite are all example of different methods to distribute and transmit books digitally (Gilmore 2007). This clause covers the various digital reading devices and technology, as well as website use.

“[...] including without limitation “e-book” format and other electronic or machine-readable media, or on-line or satellite-based data transmission.”

Subsidiary rights

The subsidiary rights deal with all the additional secondary rights, ranging from audio, to movie, serialization, foreign rights, etc. Subsidiary rights can also take digital form, and should therefore list electronic forms as part of its clause. Subsidiary rights can include electronic forms that refer to word-for-word digital versions or enhanced multimedia digital content (Gilmore, 2007). In addition, subsidiary rights should include the right to license the digital rights to third parties.

Karen Gilmore states that it is not enough to obtain rights for the publishing house; in the event a third party is involved, the contract needs to outline who can have access to these rights and under what conditions (Gilmore 2007).

Electronic subsidiary rights can affect other rights such as printed subsidiary rights and marketing rights; publishers should clarify the intended use of these rights. For example, a publisher decides to include a book in *The Globe and Mail's* online first chapter program for marketing reasons, but has licensed first serial rights to a printed magazine for the first two chapters of the book. Digital publishing opportunities cannot come into conflict with the licensing of other printed rights. Subsidiary rights should include both printed and electronic forms (Gilmore 2007).

Marketing rights

In addition to having the right to publish the work in part or in its entirety in hard copy and digitally, the publisher needs to be able to market the work using digital tools (for example by collaborating with Google and Amazon or creating websites). Although some authors might resist the creation of digital versions of their work, it is important for the publisher to minimally be able to market the work *digitally* (Davy 2007, 7).

Making available an excerpt or selection of the work, without paying royalties or compensation to the authors, is useful for marketing purposes, and this possibility needs to be specified in the marketing clause. “Without charges, without royalties” is added to the marketing rights section to obtain the permission to use selection or entire works (again, some programs demand the entire content for participation) for marketing purposes. Without this, it becomes more difficult to take part in programs like Google Book and Amazon “Search Inside the Book.” Since these programs can generate ad revenue, the publisher can add a clause mentioning that, in the case revenue is made, it will be divided between the author and publisher.

The basic marketing grant should stipulate that the publisher is granted permission to reproduce excerpts or selections in any format (that includes electronic forms and broadcasts) as well as a certain number of illustrations and/or photographs for the promotion the work (Gilmore 2007). These grants of rights give permission to the publisher to gather and create enough material to fully benefit from online marketing possibilities and from new technology.

Competing works

Karen Gilmore, during the LPG panel discussion stated that the possibility of competing works is one of the major issues with digital rights. The understanding is usually that the work submitted by the author has not yet been published. An interesting challenge arises when not everyone considers putting content online as a form of publishing. In addition, authors may want to have an excerpt or selection of their work on their own websites, which creates problems if the publisher is selling digital subsidiary rights (for example first serial rights in digital form), or plans on selling digital versions of the books. The authors and publishers need to work together to plan some of the marketing tactics, to make decisions regarding the work's digital presence, and to consider how these decisions will affect other possible revenues through digital means (Gilmore 2007).

OWNERSHIP AND PROCUREMENT OF RIGHTS

To take part in digital marketing and begin a digital strategy, publishers must have the appropriate rights. As mentioned previously, older contracts often lack some (or all) digital rights or have too vague a definition of digital rights. To remedy the situation and be more flexible, publishers must contact authors to modify and clarify contracts.

At Douglas & McIntyre, some of the titles selected for pilot projects had contracts that lacked digital rights. Karen Gilmore (Vice-President Operations, Corporate and Legal Affairs) sent a letter explaining the situation and an addendum to the contracts for authors to sign. The letter provided a detailed account of a) why the rights were necessary for the publisher to obtain, b) what the publisher intended to do if they secured those rights, and c) how the authors would benefit from the projects. At Douglas & McIntyre, the success rate was high: for Amazon's "Search Inside the Book," all authors approached signed the addendum provided by Karen Gilmore. This step has not happened yet for Douglas & McIntyre's current digital project, Best of BC Digital Reference Collection. Giving authors enough information regarding the project involving their work, and telling them *specific* areas where the rights would be *used*, demonstrated that Douglas &

McIntyre was serious about the authors' work and that the publisher was not greedy (i.e. they were not asking for all the rights without intending to use them). In other words, details of the project and of how the rights will be used are key.

Some advice sites for authors and creators recommend that authors retain their digital rights (Davy 2007, 7). The fear is that publishers will take advantage of the authors, locking them in contracts without means for the copyright to revert back to the authors. Through technology such as print-on-demand, the existence of digital versions of works mean that the work can theoretically never go out-of-print, thus preventing the return of copyright to the authors. Some authors will then prefer keeping digital rights for themselves. Publishers dealing with authors that will not license digital rights can always forgo the contract: “[The] unwillingness on the part of a creator or agent to include electronic rights can be a deal breaker” (Davy 2007, 6).

However, most authors understand that publishing is in a transition period and that technology is changing the publishing industry (Davy 2007, 7). They appreciate that publishers are in a position to benefit both the authors and themselves from these rights. Though time consuming for publishers, authors appreciate publishers first obtaining the authors' consent to use digital media. Authors' input helps the publisher avoid future problems, and might provide new or alternate ideas to exploit digital means.

CHOOSING TITLES TO BE DIGITIZED

To control their digital strategy, publishers do not simply need the rights to the titles, but ideally would have control of the digital files of their entire list (Davy 2007, 8). However, digitizing the entire publisher's list can be impractical due to time and financial constraints. Because every title does not need the same digital strategy, choosing carefully the titles to be digitized is essential. The selection of titles is dependent on both the goals and reasons for the digitization of a particular title as well as the titles currently available in different formats. Key questions publishers can use to help selecting their titles for digitization include the following:

- Will the digital files be used in more than one area of the publisher's digital strategy (such as e-books, viral marketing, book browsing, and content for websites and newsletters, etc.)?
- Is the platform of the author already part of the digital realm, or can it move to it easily? The author's platform refers to the author visibility and reputation within a particular subject, field or area, or within a particular community, which can be measured through things such as speaking engagements and events, regular columnist magazine or blogs, contacts within the community that would be interested in the author's work. In an increasingly digital world, the digital aspects of an author's platforms are becoming even more important.
- Is the subject matter particularly appropriate for online marketing, online excerpts, audio or video treatments, and/or e-books?
- If it is a backlist title, will the title be reprinted, is it selling at a continuous rate? Would a minimum marketing effort revive its sales (i.e. is the subject appropriate to today's political/cultural environment)?

In the case of Douglas & McIntyre and the pilot project Best of BC Digital Reference Collection, the selection process was made easier because of the framework for submission which was defined by target for the offer to the library consortia in BC. Titles submitted to the Best of BC Digital Reference Collection, a collection of titles from British Columbian publishers, can also be submitted to the Canadian Research Knowledge Network (CRKN) by CEL if the publishers decide to take part in both programs.⁴³ The CRKN is acquiring social sciences and humanities titles for its Canadian Electronic Library (CEL). Douglas & McIntyre could not submit their entire list, because they had to choose titles that would be appropriate for the CEL, that is titles that fit the criterion of the project, specifically non-fiction titles within the subjects of social sciences and humanities. Having a pilot project, whether an internal project or one from an external source,

⁴³ CRKN recently received a federal grant from the Canada Foundation for Innovation (CFI) to make social sciences and humanities titles available in digital format, thus making research and information widely available for researchers and students..

but with parameters clearly defined, can create the necessary structure for a publisher to begin the digitization of titles, and to slowly begin to test the possibilities of digitized content for publishing.

At Douglas & McIntyre, the process started with a preliminary selection of titles that fit the pilot project's restrictions. Then all contracts from this first selection were reviewed to verify which digital rights, if any, were included: the Douglas & McIntyre interns looked at each of the pre-selected titles' contracts to determine if they included the old or new versions of the primary rights, subsidiary rights, and marketing rights.⁴⁴ The next step was to establish the digital status of the files: which titles already had what kind of digital files available, and which titles needed to be converted into digital versions (which titles had PDFs of the entire book, which titles only had digital versions of the covers, which titles had digital files in layout programs such as InDesign and QuarkXpress, and which titles only existed as hard copies).

Most recent titles are already available in digital formats, often as PDFs, JPEGs, and/or original digital layout files (most likely produced in InDesign or QuarkXpress), and as such these titles are already digitized. For titles with digital files that need to be converted, this conversion of older digital files into new ones can be done in-house by production staff. At Douglas & McIntyre, once the digital status of each title was determined, where possible an employee converted the files to the newest PDF format. Because it can be done in-house, it means that all recent⁴⁵ titles can be easily digitized (assuming all rights needed are obtained) and it permits publishers to maintain complete control of their digital content.

However, the digitization of the backlist titles is more complex: it takes more time to do so and may necessitate the assistance of a third party.⁴⁶ Since digitizing the backlist is more complex

44 Once publishers complete the analysis of the types of digital rights (and acquire the rights missing for digitization), publishers can use these rights for future digital projects without investing again the time to analyze the contracts.

45 "Recent titles" refer to frontlist titles and backlist titles that have some form of digital files available.

46 For more information on files formats and conversions, please refer to section "What does digitization look like?"

and more time-consuming, selecting the appropriate backlist titles is important. At Douglas & McIntyre, the selection process of the backlist is based on the rights held, the different uses for the digital content of the works, and the parameters of the pilot project, the Best of Best Digital Library Collection.⁴⁷ Karen Gilmore describes the selection process as such: a) If the title is still “in-print,” but Douglas & McIntyre currently does not have many physical copies in stock, the title may not benefit from entering Amazon’s “Search Inside the Book” and Google Book Search programs, especially if the respective author contract does not include digital rights and the title needs to be scanned. b) However, this same title may be digitized if there is potential interest in that book in digital form (e-book), or if including it in different lists, such as the Best of BC Digital Reference Collection, adds authority and value to the collection. In this case, as the interest grows for a digital copy of the works, the title may then be submitted to Amazon and Google at a later time. The more the digital content of one title can be used for different aspects, the more valuable the digitization process is for that title. c) Finally, part of the consideration in the selection process is the cost of digitization. The quality and therefore the flexibility of the digital files will vary according to cost. For example, a simple scan will cost less than a text files extracted from a scan through OCR software. In other words, the purpose of the digital files will also affect the cost: creating a digital version for Amazon “Search Inside the Book,” Google Books, or CEL is less complicated and less expensive than the creation of files for the formats necessary for some e-readers or for print-on-demand (as they demand more information about the files, more flexibility, in-depth indexing and extensive metadata).

Another important consideration for digitizing titles is the type of agreement the publisher has with the distributors of content. At Douglas & McIntyre, in the context of the pilot project, the publishing house specified that they would license the titles under a perpetual access license for multi-users and be non-exclusive. The non-exclusivity agreement permits Douglas & McIntyre

47 The parameters are non-fiction titles within the subject areas of humanities and social sciences.

to make their titles available through other distributors. Douglas & McIntyre can then make their digital versions of their titles available to various distributors as they see fit.

However, Karen Gilmore explains that the perpetual access for multi-users means that the title selection for the collection needs to be done carefully: the on-going access from libraries and universities by multi-users simultaneously can mean that a class of students could all forgo purchasing a title by accessing it online through their university. For titles that reach few students, this is a non-issue, but for a title that may become a standard required reading, it may result in lost sales for the publisher. The publisher needs to contemplate each title's potential before including them in projects with different licensing agreements.

When selecting titles to be digitized that will specifically be used for online marketing (for a new title or to revive a backlist title, which are already digitized), publishers need to pay attention to authors' platforms. In the event that a title is perfect for a pilot project and that it could take part in various established digital programs, the publisher next considers the performance of the author platform *online*.⁴⁸ Authors who already have a strong presence online are good choices to test online marketing, whereas authors who do not appeal to online or tech-savvy audiences might make online marketing efforts less beneficial. Both the current online presence and willingness of authors are important. Although not all books will benefit from a strong online platform, if only a selection of titles from a publisher's list will be digitized, books that benefit from online marketing tend to be books where the author's platform already has an online presence or can easily be inserted on the web.

Finally, publishers need to incorporate digitization process within their acquisition practices. When signing contracts, the different digital rights need to be a standard part of the contract and the title's digital strategy needs to be considered throughout the brainstorming and concept decisions (from marketing, to subsidiary rights, etc). This decision process becomes increasingly

48 The author's platform is always important, but specific attention should be paid to the web spaces if the title will be promoted online.

complex as the technology rapidly changes. Within the usual two-year span between the contract signing and the publication date, a lot can change in the digital world.

SECTION CONCLUSION

The selection process for digitization can be deciding to digitize the entire list or can be a more complex process where the publisher considers what digital opportunities lie ahead for each individual title. Yet, the most important aspect of digitization centers on having the correct digital rights to begin a digital strategy. Publishers must consider the different opportunities for the titles, the licensing agreements with the different digital file distributors, and the authors' online platforms. The selection process should not only take place at the end of the production of a title, but should begin in the conceptual phase of every new book. Digitization opportunities can be discussed at the very beginning of the publishing process.

WHAT DOES DIGITIZATION LOOK LIKE? (FORMATS, PROCESSES AND MANAGEMENT OF DIGITIZATION)

INTRODUCTION

One of the major difficulties when embarking on the digital road comes from the number of different file formats, platforms, processes, and third-party service providers available to publishers. Once the decision to digitize content and the selection of titles has been made, the next step is to identify and select the methods to create and use the digital files. This section explores the choices available to publishers in terms of formats, processes, and service providers that can assist with the digitization process. In addition, this section addresses the issue of standardization as a necessary goal for the publishing industry to better streamline and simplify the digitization process.

FORMATS

Since there are no set standards for formats in the publishing industry or clear preferences on the part of users, publishers face the hard decision of either choosing specific platforms (and therefore business partners) or attempting to accommodate most platforms by creating multiple versions of each title. Although there are many file formats for digital content, formats can be broken down into two main categories: flowable text and paginated text (with the possibility of combining the two and using them with other programs such as Java and Flash). The main difference between the two formats resides in their degree of flexibility to control and manipulate text.

Flowable text

Flowable text refers to a format common on the web. This format is extremely useful because it can be “flown” from one platform to another since the text easily adapts to its new platform’s structure, and because it can reformat itself on a singular platform. For example, the text refor-

mats itself when the user of an e-reader increases or decreases the size of the type. Instead of simply zooming in and out of a page, flowable text simply rearranges itself to fit the dimensions of the platform. File formats that use flowable text are based on text and markup (Knight 2007, 7). Markup language is the added information that dictates how the text is to be formatted on the target platforms; for example, it can specify the formatting of the text for both a PDA and for a website. The text will look different (different line length, fonts, leading), but the structure of the text will remain the same; for example, the hierarchy of the text will be the same on either platform. Flowable text is markup language, such as HTML (HyperText Markup Language), XHTML (eXtensible HyperText Markup Language) and XML (eXtensible Markup Language).

Flowable text permits great flexibility for changing the presentation of the text: the content can break away from the traditional definition of the book as an object. Content, no longer restrained by pages, covers or linear reading, is useful since it can be reused or repurposed quickly and efficiently into different formats for different uses (for example, the content as flowable text can be used in book form and on a website without spending too much time formatting the text on each platform). But, flowable text also threatens publishers: revisiting the definition of the book (since the content is no longer restrained by the book format) leads to questions on the role of publishers. Publishers who see books as vehicles for content (for example, mass market titles) can see flowable content as an interesting addition for the distribution of their content. However, for publishers who attend to the book format, such as design details, flowable content that breaks the codex barrier, flowable text is not pragmatic.

Paginated text

Whereas flowable content is successful thanks to its flexibility, paginated text presents the opposite advantage. Paginated content provides more control over the text and, in turn, over the files and how they are used. Paginated content, most common as Adobe PDF, maintains the format chosen by the publisher. It is an accurate digital version of the original book, where the structure, page

size, fonts, and leading remain the same; in other words, the text is not flowable (Knight 2007, 9). Nonetheless, paginated text is not completely inflexible. Paginated content can contain metadata, external and internal links, and embedded multimedia files (such as videos and audio files) inserted into the paginated file. The flexibility of paginated text does not come from the text, but from the publisher's control of how the content is presented and of the restrictions applied to the use of the digital files.

Not all paginated formats are created equally: the quality of the file depends on its flexibility and accessibility. Most e-book distributors ask that publishers give them PDFs that have the following characteristics (Gibson 2006; codeMantra, 2005, 6):

- all fonts are properly embedded
- security removed (each e-book distributor can then apply their own DRM measures)
- table of contents are bookmarked and tagged (this is for better usability: users can simply click on the table of content to access the desired page)
- links (internal and external/hyperlinks) are turned on (again for enhanced usability: users can simply click on links to access internal references or external links on websites)
- endnotes are tagged (rather than search manually through the document, users can have easy access to endnotes if they are tagged)
- files are compressed
- all printer's marks are removed.

PROCESSES TO CREATE DIGITAL CONTENT

The flexibility of the files, in terms of form and not just content, is important. The flexibility and quality of the files greatly depend on the state of the publisher's existing digital files. The technology available at the time the title was published will affect the title's flexibility. For recent titles with existing digital files, the conversion to the various formats is simple. However, unless the publishing company has been in existence for less than fifteen years, an enormous number of

titles do not fall into this category. For older titles there are four processes of digitization: image scanning, extracting text from scanned images with OCR software, tagging (of markup language) and including metadata information to documents.

Image scanning

Scanning produces an exact image of a page. Scanned files are typically JPEGs or PDFs. This is the simplest and cheapest digitization process, and represents the minimum a publisher can do to create digital content from a *printed* source. The scanning can be done in-house with an inexpensive scanner or can be done by a third party with high-capacity equipment that delivers high-resolution files at a very efficient speed. Scanning can be done with non-destructive, keeping the book intact, or destructive techniques, cutting the spine and feeding the pages through the scanner. The destructive process delivers files faster and with greater accuracy (LibreDigital 2007, 26-27); by outsourcing the scanning process to large companies, books can be scanned at 1,000 pages per hour without the assistance of a human page-turner (Kelly 2006).

Creating image files of books through scanning makes it possible for publishers to at least make the content visually available online.⁴⁹

OCR (Optical Character Recognition)

Once the book has been scanned, a publisher can extract text from the image file by using a Optical Character Recognition (OCR) software. As for scanning, OCR can be done in-house by purchasing the software necessary and meticulously editing the new digital text against the original book. Alternatively, third-party providers can extract the text from image files. The accuracy of the new text will depend on editing. Accuracy can increase by using more than one software,⁵⁰

49 However, publishers will find that this solution creates a problem of big file sizes.

50 One can use two or three different brands of OCR software to extract the text and then compare and compile the results into one document.

using various dictionaries, and/or having a copy-editor manually verify the final text (LibreDigital 2007, 27). This digitization process is important to make the content digitally available on platforms that are dependent on flowable content, for example to place excerpts on websites and to create searchable content.

Tagging

Once raw text has been created, it can then be tagged for formatting purposes. The text is tagged with additional information that specifies how the text should be formatted. This level of tagging is important if the book's content will be “flowable.”

Metadata

Metadata is applied once the text has been formatted into the desired book layout and a PDF has been created (i.e. the content is now paginated text) (this process can also be referred to as “tagging”). The PDF can then be tagged with additional information about the file and content (i.e. metadata). There are many levels of metadata: it ranges from general information about the copyright and the publisher to precise line-level information, depending on the use of the content. This is important in terms of the usability of e-books or digital content.

Image scanning, the creation of flowable text through OCR software, and adding metadata to PDFs are all examples of decisions to be made in the creation of digital files of the publisher's titles.

MANAGING AND DISTRIBUTING DIGITAL ASSETS

DAMs and DADs

Once content has been digitized, publishers need to manage and distribute their digital assets. A new category of service providers has entered the publishing industry to help publishers meet

their needs for digital asset management and distribution. These service providers, known as DAMDs⁵¹ are contracted to implement the production aspects of the digital strategy; they relieve the staff from having to deal with the production, management and distribution of digital content. Essentially DAMDs are the intermediary between the publisher and e-book distributors.

DAMDs⁵² offer services like e-book creation, composition and typesetting, technical edition and copyediting, and image processing. However, their primary and most useful service is data conversion. DAMDs first convert publishers' titles from print-ready files or hard copies into files ready for distribution, and second, ensure that files are compatible with requirements of major e-book distributors, such as NetLibrary, ebrary, OverDrive, Lightning Source, Adobe, Amazon, and Google (codeMantra 2005, 5). Unlike Google, who essentially also displays digital content, DAMDs "offer a multi-faceted distribution 'solution' for publishers" (Shatzkin 2007b). In other words, DAMDs are not simply responsible for the distribution of content in limited digital formats to limited platforms, but must attempt to deliver a complete wide-scale digital distribution service (i.e. delivering digital content to any platform and website in any desired digital format) (Shatzkin 2007b). DAMDs also process, store and deliver content, and apply DRM measures (LibreDigital 2007, 4). DAMDs are an appealing option for publishers who prefer not having to deal with quickly converting each title to accommodate various e-book platforms (Shatzkin 2007b). They ensure that all metadata required is included, and they guarantee that the file standards are met. Providers of DAMD services retained by a publisher should guarantee an extensible system: one that can grow and accommodate change in the publishing house and in digital technology, one that can incorporate flexible rights and permissions for access to the system, and one that reflects the publisher's need for management and metadata input (LibreDigital 2007, 6; Klopotek 2007, 3).

51 The common terms for digital asset managers and distributors are DAMs and DADs, but DAMDs refer to companies that offer both management and distribution services for digital assets (Knight 2007, 5).

52 Some emerging DAMDs are codeMantra, BiblioVault (University of Chicago), Accenture, Donnelley, LibreDigital (HarperCollins and NewStand), Holtzbrinck (BookStore), and Ingram Digital Ventures (Shatzkin, 2006).

In-house versus third-party service providers

The most beneficial aspect of DAMD is the outsourcing of the conversion of data for maximum compatibility with different e-book distributors. Although time consuming, this particular task can be accomplished in-house. Most e-book distributors have similar digital file requirements: files that have embedded fonts; a three-level linked and bookmarked table of contents; linked endnotes; images optimized for screen, image and text disclaimers; no printers' marks; and proper pagination (Gibson 2006, 16). These requirements can be met in-house as part of the production process. Keeping the digitization process of titles in-house gives publishers the opportunity to be more flexible with their overall digital strategy. Large publishing companies, such as HarperCollins, have started developing their own warehousing and managing systems for digital content to avoid having to depend on third-party providers.⁵³ Small-to-mid-sized publishers do not need the same extensive systems and can use software available to help manage their digital assets.⁵⁴

When deciding to keep titles in-house or use the services of a third party, publishers should consider the following questions.⁵⁵

53 HarperCollins invested in NewStand, who built HarperCollins' digital warehousing system LibreDigital (Milliot, 2007). Other publishers can use HarperCollins' services: the digital conversion service is directly with HarperCollins while the warehousing and distribution is with LibreDigital (Shatzkin 2007b).

54 Management software such as Canto (<http://www.canto.com>), North Plains (<http://www.northplains.com>), and Artesia (www.artesia.com). These software are created to archive and manage different digital assets used during the publishing process. The software boast that they are easy to install and use, thus enabling the customers to create a low-cost digital asset management solution for their company.

55 These following questions do not address cost, however to provide an approximation of DAMDs costs, please refer to Appendix C for codeMantra's breakdown.

Who will or needs to control the assets?

DAMs are most useful for the conversion of digital content into flexible formats for e-book distributors. A publisher with limited revenue from e-book sales but who uses diverse digital files and formats for other ends (such as marketing) may need greater control of the digital assets and may prefer to keep all assets in-house. However, for publishers dependent on sales to large end-users like libraries and universities, who acquire substantial numbers of e-books, DAMs can present the advantage of simplifying the aggregation and sale of their digital content.

Does the publishing house currently have a high level of external/internal exchanges of digital material?

A publishing house with high traffic in digital files⁵⁶ would benefit from a centralized and automated system granting rights and permissions to the people involved without the intermediary of an employee controlling the constant digital traffic. A system provided by a DAMD or through the purchase of management software can simplify digital file sharing. This is especially important for larger organizations or companies that depend heavily of freelancers/contractors. However, for largely in-house publishing groups with low exchanges of digital content, simply using the company's server, FTP site, or email, is sufficient to manage incoming and outgoing digital files (with proper organization on their part).

Does the publishing house own a large number of digital assets that can no longer be handled by a simple organized file server?

An organized server where all the files are consistently labeled and saved according to previously determined guidelines is sufficient to begin the digitization of titles. In order to organize files optimally, for example, all files related to one book start with the ISBN number. However, once the

56 Meaning many files are exchanged digitally either internally within the publishing house or externally, such as with other publishers, publicists, printers, etc.

number of digital files increases to the point where managing files becomes too complex, or the server can no longer sustain it, an automated system is needed.⁵⁷ The automated system can keep track when a new file is uploaded and whether the correct file formats are uploaded, and it can send automated messages to individuals notifying them of new uploads by different members of the system (Knight 2007, 44-46).⁵⁸

Does the publishing house already have a system that incorporates metadata?

Metadata is key for digital assets. It preserves information about the content and the file, it allows search capacities, and it works with DRM measures to control access and use of the content. Having the capacity to add and manage metadata information for every title is crucial. Publishers with access to systems such as Acumen already have the possibility to include metadata, for each work (Acumen is a repository of metadata and therefore manages it; it does not add metadata to the digital files of titles). At Douglas & McIntyre, the Acumen system allows the editor to input the metadata in several fields such as the title, author(s) and contributor(s), ISBN, publication date, size, format, language, Canadian and American prices, and edition. This information, the metadata, can be used to create ONIX (Online Information Exchange) reports or generate excel spreadsheets based on particular metadata searches (for example, during the creation of a new Douglas & McIntyre website, information on all titles from the past six seasons needed to be retrieved. Through a search based on metadata (i.e. the publication date in this instance), the user of Acumen was able to quickly access these records without having to sift through each individual

57 This is left to the publisher to decide when it is appropriate, but a good time to make the shift is if the file server keeps crashing or when the in-house staff responsible can no longer attend to the rest of their workload.

58 Alison Knight reports in detail the codeMantra process at UBC Press in her Master of Publishing project report *The Tangled Web*. For example: "[A] cover designer submits completed cover designs directly to Collection Point [the codeMantra "hub" for file sharing between parties involved in the titles' creation], including printer instructions if required. When these uploads occur, the title's production editor is automatically notified by email. By using uploads and local downloads through Collection Point [...] all parties involved in the production of a title ensure that they are using the correct and most recent file" (Knight 2007, 46).

title. Publishers without similar systems should look into either obtaining one or using a DAMD that provides metadata incorporation among its services.

Does the publishing house anticipate updating layout software, such as InDesign CS3, that can produce e-book files?

Increasingly, people in the publishing industry are striving for digital standards and streamlined production processes. Adobe InDesign CS3 gives the ability for in-house staff to create e-book files (EPUB format) directly from an InDesign file, simplifying the e-book conversion process. For publishers using recent Adobe software, the need for third-party providers may be greatly reduced since they can create the appropriate e-book files themselves.

Does the publishing house have the opportunity to take part in pilot projects or consortium that may decrease costs for conversion, management and/or storage?

Some pilot projects currently underway, for example the Best of BC Digital Reference Collection with Gibson Publishing Connections (GPC), offer publishers reduced costs for the conversion, management, and storage of digital assets, making the process simple and cost effective. In the case of the Best of BC Digital Reference Collection, GPC offered to pay all conversion costs for up to forty titles selected,⁵⁹ and the cost would then be deducted from the royalties generated from the sale to libraries (in this case, to CRKN) (Gibson Publishing Connections).⁶⁰ Therefore only the digital files generating profit required the publisher to pay (and only through the deduction from royalties earned) for the conversion of the digital files. Copies of files converted by GPC are to be

59 The titles selected refer to the titles submitted by the publisher and then accepted by GPC to be included in CEL/Best of BC Digital Reference Collection.

60 GPC benefits, in this case by establishing itself as a reliable DAMD and hoping to generate future business. For the actual GPC pricing, please refer to Appendix C.

given to the publishers for their own uses (Gibson Publishing Connections).⁶¹ In today's digital environment, pilot projects are probably one of the best opportunities for small trade publishers to venture with digitization. In addition to new developing digital pilot projects for publishers, many publishers received funding from the Canada Council for their digital projects. The funding was awarded to publishers who submitted project proposals that dealt with digitization or online initiatives (Grittani-Livingston 2007b).

STANDARDIZATION

The lack of format and platform standards in digital publishing is one of the greatest obstacles for digitization of titles. This lack also created the need for DAMDs for publishers: if there were standards for digital formats, publishers would not need the services of DAMDs since they primarily provide the efficiency of converting digital files to all the different formats needed to fulfill the needs of the different platforms. If digitization of content can be efficiently incorporated in-house as part of the workflow, publishers can keep control of the files and content more easily. Attempts at standardization have been made most recently by the International Digital Publishing Forum (IDPF). Its mission is to create standards for the electronic publishing industry to benefit readers/users, publishers, authors of content, and platform developers (www.idpf.org). The forum members believe that open standards, which are free and facilitate the exchange and use of information across platform and formats, are required to foster a sustainable digital publishing practice. Hence they developed the EPUB, a XML-based flowable file format. Ideally, all e-book platforms will support this EPUB format:

[EPUB] allows publishers to produce and send a single digital publication file through distribution and offers consumers interoperability between software/

61 GPC is unlike Google or Amazon in that they converted files only for their benefits and did not share the digital files with the publishers).

hardware for unencrypted reflowable digital books and other publications (www.idpf.org).

Nick Bogaty, founder of IDPF, believes that standardization of the format of e-books would reduce e-book-production cost and turnaround time all the while remaining easily available to all publishers; it would also represent an increase of titles available to readers of digital content (Bogaty 2007, 6).

Publishers interested in standardization of e-books need not understand the intricate details of standards. Adobe, with its newest program (CS3), has created an authoring tool for publishers to be able to create their own e-book files in-house and to wrap files in protective DRM technology. Adobe CS3 supports the creation of EPUBs (as an export function) and the OPS Standards (Knight 2007, 59). For OPF Standards, companies such as Adobe, e-book Technologies Inc., OSoft, VitalSource Technologies, Sony and LibreDigital have all embraced and implemented these standards for their e-readers. Standardization of e-books will be possible only if both publishers and e-readers creators follow established standards.⁶²

SECTION CONCLUSION

After deciding to develop a digital strategy, choosing the appropriate titles for digitization, and acquiring all necessary rights for the digital strategy, publishers must decide ways in which the titles will be digitized, and ways the digital assets will be managed and distributed. This section addressed these last aspects of the digitization process. First, the section addressed the different formats available for digital content (flowable and paginated content). Second, the section explored the different processes to create digital files (scanning, OCR, tagging, and adding

62 Hachette Book Group USA recently announced that it will use the EPUB file format for its e-books, supporting file format standardization. It remains to be seen if many other publishers will follow in their footsteps, especially if it is now easier to create EPUB files through Adobe CS3. Unfortunately, Amazon's new e-reader "Kindle" does not support EPUB.

metadata). Third, the section discussed the different alternatives for managing digital assets. The section concluded with a discussion of standardization of format to benefit the publishing industry and the users.

C O N C L U S I O N

Digitization should not be feared by small-to-mid-sized Canadian trade publishers to take advantage of new opportunities offered by the digital world. Creating a digital strategy is not as complex as it might first seem. Once the decision to develop a digital strategy is made, efforts to incorporate it within the general publishing process are necessary if the strategy is to be useful. It has been established that for publishers to keep greater control over the digital assets, the digitization process should remain in-house as much as possible. However, for a truly successful digital strategy, the digitization of titles should be considered throughout the publishing process:

- *Concept/editorial:* What is the digital platform of the author(s)? How does the subject translate into online spaces? Who will be responsible for metadata input? What are the best formats for author-publisher exchanges?
- *Rights:* What are the digital rights necessary for a title's digital strategy? What subsidiary digital rights can be licensed?
- *Marketing:* What are the most appropriate online marketing and promotional opportunities?
- *Production/design:* How will the files be converted into the appropriate and necessary formats? How and when is the metadata going to be included? How and who will be responsible for DRM measures?
- *Distribution:* How will digital content be distributed?
- *Sales:* Who are the online retailers? Will the publisher's own website be used for digital and/or hard copy sales?

Incorporating the digital process in publishers' current workflow is a necessary step. Streamlining the process of digitization gives publishers the possibility to further exploit their digital assets. As demonstrated in this guide, there are multiple opportunities for the digitization of the book publishing industry.

Through the course of my internship at Douglas & McIntyre, I witnessed the early stages of their digital strategy's development. The pilot project Best of BC Digital Reference Collection was a unique opportunity for Douglas & McIntyre. Through this project, they were able to dedicate resources to establish the titles that should and are going to be digitized, and to explore possibilities for these titles other than that of being part of the Canadian Electronic Library (CEL). They were able to verify the status of the digital rights of their author contracts, analyze the different methods to digitize content, identify some of the service providers of digitization, and make decisions about the storage and maintenance of digital assets and rights. Since this pilot project is still underway, and some of the decisions are still pending, no final verdict has been made regarding the project and the final digital strategy. Yet, it is clear that this project was the impetus to the development of a sound digital strategy. This strategy is not limited to one project; Douglas & McIntyre has looked into various possibilities that digital technology offers publishers. With their digital strategy, Douglas & McIntyre will be able to better incorporate digital aspects into their publishing process, giving them the flexibility to both take part in current digital opportunities and also be ready for future digital changes.

Going Digital: A Guide for Book Publishers was inspired by Douglas & McIntyre's ongoing development of their digital strategy. This guide aimed to demystify the process and encourage similar publishers to follow suit. The goal of this guide is to bring a greater understanding of the evolving digital practice in the publishing industry. As there are yet no rules or fixed standards, publishers must simultaneously prepare themselves for an uncertain digital future as well as exploit present digital opportunities. This guide aims to give publishers as much knowledge as possible about today's digital trends. Staying abreast of changes will be important for publishers to stay current, and hopefully as well as to take part in influencing the practice of digitization for the better.

The guide was divided into four sections. The first one, "Context" briefly explored the possibilities and constraints associated with today's changing technological environment. It also

discussed one of the most debated issues of publishing and technology: copyright issues dealing with content available digitally for free or for a fee.

The subsequent three sections of the guide covered the process of establishing a digital strategy. The section “Why go digital?” outlined the different digital opportunities available to publishers: websites, viral marketing, book browsing and searching, and e-books, were listed as integral part of the motivations for digitization.

The section “What do you need” explored the differences and subtleties of digital rights. To pursue their digital strategies, publishers need to obtain the different digital rights (primary, subsidiary, and marketing right). The section also briefly described the process of obtaining digital rights for older contracts. Finally, this section covered the selection process for the titles to be digitized (especially necessary for older titles that do not already have some form of digital file available).

Finally, the section “What does digitization look like?” covered the different file formats currently available for the digitization of books, and the different processes to digitize. It explored issues regarding the management, control, and distribution of the digital content. And finally, it discussed goals of digital standardization within the publishing industry.

This guide aimed to answer common questions publishers currently have about the digitization process in order to provide information to small-to-mid-sized Canadian publishers developing or wishing to create their own digital strategies. This guide is meant to inspire and encourage publishers to take control of their titles as the trade publishing industry faces transition brought on by the impact of the digital age: creating a digital strategy provides a means for publishers to not only be prepared for an uncertain future by being flexible, but also answers current needs, more specifically for online marketing. Every publisher needs to contemplate how digitization will affect their practice and how a digital strategy can benefit them and their titles.

B I B L I O G R A P H Y

- Amazon. URL: <http://www.amazon.com> (accessed 12 October, 2007).
- Associated Press. "Newspapers Hope for Online Growth." *Globe and Mail*, 5 December, 2007. URL: <http://www.theglobeandmail.com/servlet/story/RTGAM.20071205.wgtnews1205/BNStory/Technology/home> (accessed 6 December, 2007).
- Beschizza, Rob. "Killed By DRM: e-Books." *Wired*, 26 April, 2007. URL: http://blog.wired.com/gadgets/2007/04/killed_by_drm_e.html (accessed 6 December, 2007).
- Bogaty, Nick. "Digital Books: Why This Time It's For Real." Presentation, Book Expo America, New York City, NY, 31 May, 2007.
- CBC, "HarperCollins Moves to Control Digital Rights." 13 December, 2005. URL: <http://www.cbc.ca/story/arts/national/2005/12/13/digital-books.html> (accessed 5 November, 2007).
- codeMantra and Gibson Publishing Connections. *Digital Content Distributions with Universal PDF (SM): A White Paper*. Plymouth Meeting, PA: codeMantra, 2005.
- Crawley, Devin. "E-books Face a Tough Road in Canada." *Quill & Quire Omni*, 1 May, 2001. URL: http://www.quillandquire.com/news/article.cfm?article_id=2032 (accessed 6 December, 2007).
- Davy, Diane. *The Impact of Digitization on the Book Industry: A Report Prepared for the Association of Canadian Publishers*. Toronto, ON: Castledale Inc., 2007.
- Deahl, Rachel. "HC Launches Browse Inside Feature." *Publishers Weekly*, 26 February, 2007. URL: <http://www.publishersweekly.com/article/CA6419369.html?q=HC+widget> (accessed 14 December, 2007).
- Drassinower, Abraham. "Taking User Rights Seriously." In *In the Public Interest: The Future of Canadian Copyright Law* by Michael Geist, 462–479. Toronto, ON: Irwin Law, 2005. URL: <http://www.irwinlaw.com/books.aspx?bookid=120> (accessed 22 January, 2008).
- e-books.com. URL: <http://www.e-books.com> (accessed 12 October, 2007).

- Gibson, Bob. *Canadian Electronic Library: Monographs from Canadian Publishers Organized for the Digital Library*. Saint-Lazare, QC: Gibson Library Connections Inc., 2007. URL: <http://www.gibsonlibraryconnections.ca> (accessed 14 July, 2007).
- . *Digital Asset Management and File Conversion for Canadian Publisher*. Saint-Lazare, QC: Gibson Library Connections Inc., 2005. URL: <http://www.gibsonlibraryconnections.ca> (accessed 14 July, 2007).
- . *Publishing for the Digital Library*. Presentation for the Association of Canadian Publishers. Gibson Publishing Connections, 6 June, 2006. URL: http://www.publishers.ca/Publishing_for_the_Digital_Library.pdf (accessed 14 October, 2007).
- Gilmore, Karen. "Contract Issues in the Digital Age." Literary Press Group panel discussion *Rights and Contracts in the Digital Age*, Vancouver, BC, 4 May, 2007.
- Google. *Google Books Search: The Story*. Information booklet and package for promotion, London Book Fair, UK. URL: <http://books.google.com> (accessed 11 October, 2007).
- Grittani-Livingston, Megan. "Random House of Canada Makes Books Searchable Online." *Quill & Quire Omni*, 25 May, 2007a. URL: http://www.quillandquire.com/news/article.cfm?article_id=7770 (accessed 6 December, 2007).
- . "Publishers Go Digital." *Quill & Quire Omni*, 1 July, 2007b. URL: http://www.quillandquire.com/news/article.cfm?article_id=7909 (accessed 12 December, 2007).
- HarperCollins, "e-books." URL: <http://www.harpercollins.com/imprints/index.aspx?imprintid=517980> (accessed 12 October, 2007).
- iCopyright. URL: <http://www.info.icopyright.com> (accessed 25 June, 2007).
- IDPF. URL: <http://www.idpf.org> (accessed 6 October, 2007).
- . *Open Publication Structure (OPS) 2.0 v0.9847* (Draft Document). New York City, NY: International Digital Publishing Forum, 11 July, 2007. URL: http://www.idpf.org/2007/ops/OPS_2.0_0.987_draft.html (accessed 6 October, 2007).

- Johnson, Tim. "Marketing by the Facebook." *Quill & Quire Omni*, 1 September, 2007a. URL: http://www.quillandquire.com/news/article.cfm?article_id=7938 (accessed 14 December, 2007).
- . "Self-starters." *Quill & Quire Omni*, 1 September, 2007b. URL: http://www.quillandquire.com/news/article.cfm?article_id=7618 (accessed 14 December, 2007).
- Kaplan-Myrth, Andy. "Creative Commons Canada." Literary Press Group panel discussion, *Rights and Contracts in the Digital Age*, Vancouver, BC, 4 May, 2007.
- Kelly, Kevin. "Scan This Book!" *The New York Times*, 14 May, 2006. URL: <http://www.nytimes.com> (accessed 13 June, 2006).
- Kerr, Ian R. "If Left to Their Own Devices...: How DRM and Anti-circumvention Laws Can Be Used to Hack Privacy." In *In the Public Interest: The Future of Canadian Copyright Law* by Michael Geist, 167-210. Toronto, ON: Irwin Law, 2005. URL: <http://www.irwinlaw.com/books.aspx?bookid=120> (accessed 22 January, 2008).
- Klopotek. "Why Every Publisher Needs a Digital Asset Distributor (DAD)." In *Digital Asset Distribution for Book Publishers: An Emerging Infrastructure*. New York, NY: Klopotek North America, Inc. 21 June, 2007. URL: <http://www.klopotek.com> (accessed 22 January, 2008).
- Knight, Alison. *The Tangled Web: Managing and Confronting Scholarly E-book Production at UBC Press*. MPub Project Report, Simon Fraser University, 2007.
- LibreDigital. *A Digital Warehouse for Book Publishers: A LibreDigital White Paper Version 3.1*. Austin, TX: LibreDigital, 2007.
- Lightbox JS. URL: <http://www.huddletogether.com/projects/lightbox2> (accessed 25 November, 2007).
- Lightning Source Inc. URL: <https://www.lightningsource.com> (accessed 12 October, 2007).
- Macfarlane, Amy. "Raincoast Podcast Attracts Hundreds in First Weekend." *Quill & Quire Omni*, 22 November, 2005. URL: http://www.quillandquire.com/news/article.cfm?article_id=6974 (accessed 14 December, 2007).

- Mandel, Charles. "Better Marketing Through Blogs." *Quill & Quire Omni*, 1 April, 2005.
URL: http://www.quillandquire.com/news/article.cfm?article_id=6626 (accessed 14 December, 2007).
- Maughan, Shannon. "Way Cool: Marketing and the Internet." *Publishers Weekly*, 2 February, 2007.
URL: <http://www.publishersweekly.com/article/CA6417182.html?q=online+marketing> (accessed 14 December, 2007).
- Maul, Kimberly. "Barnes & Noble Works with Zinio for 'See Inside.'" *The Book Standard*, 16 October, 2007. URL: <http://www.the-bookstandard.com> (accessed 16 October, 2007).
- Moynihan, David. "Adobe Captures e-book Standard. What Now?" *The Register*, 15 September, 2007. URL: <http://www.theregister.co.uk/2007/09/15/e-book> (accessed 6 October, 2007).
- Nitterhouse, Denise. *Digital Production Strategies for Scholarly Publishers*. Chicago, IL: University of Chicago, 2005.
- O'Donnell, Mike. "Email This, Print This, Steal This." *iCopyright News*, 15 March, 2006a. URL: <http://info.icopyright.com> (accessed 22 January, 2008).
———. "iCpoiright's Advertiser-Supporter Free Uses." *iCopyright News*, 17 March, 2006b. URL: <http://info.icopyright.com> (accessed 22 January, 2008).
- Offermann, Stefan. "Adobe Digital Edition 1.0 Now Available." *Adobe Press Release*, 19 June, 2007.
- O'Reilly Media, "Academic Solutions." URL: <http://academic.oreilly.com> (accessed 12 October, 2007).
- OverDrive, Inc. URL: <http://www.overdrive.com> (accessed 25 June, 2007).
- Perry, Mark. "Rights Management Information." In *In the Public Interest: The Future of Canadian Copyright Law* by Michael Geist, 251-266. Toronto, ON: Irwin Law, 2005. URL: <http://www.irwinlaw.com/books.aspx?bookid=120> (accessed 22 January, 2008).
- Project Gutenberg. URL: <http://www.gutenberg.org>; URL: http://www.gutenberg.org/wiki/Gutenberg:The_History_and_Philosophy_of_Project_Gutenberg_by_Michael_Hart (accessed 23 November, 2007).

- Publishers Weekly. "Avon Adds Video to e-Books; More Original Erotica Coming." *Publishers Weekly*, 20 August, 2007a. URL: <http://www.publishersweekly.com/article/CA6470379.html?industryid=47152> (accessed 10 December, 2007).
- . "Bloomsbury Offers New Online Book Preview." *Publishers Weekly*, 17 April 2007b. URL: <http://www.publishersweekly.com/article/CA6434398.html?q=amazon+search+inside> (accessed 14 December, 2007).
- PW NewsLine. "Amazon: Search Feature Increases Sales, Statistically." *Publishers Weekly*, 31 October, 2003 URL: <http://www.publishersweekly.com/article/CA332977.html?q=amazon+search+inside> (accessed 14 December, 2007).
- Reid, Calvin. "Amazon.com Unveils New e-Book Reader." *Publishers Weekly*, 19 November, 2007a. URL: <http://www.publishersweekly.com/article/CA6504049.html?q=e%2Dbook+ink> (accessed 14 December, 2007).
- . "Google, Widgets Mark IDPF Afternoon" *PW Daily*, 9 May, 2007b. URL: <http://www.publishersweekly.com/article/CA6440903.html?q=widget> (accessed 5 November, 2007).
- Riley, Duncan. "In Japan Half the Top Selling Books are Written on Mobile Phones." *TechCrunch*, 2 December, 2007. URL: <http://www.techcrunch.com/2007/12/02/in-japan-half-the-top-selling-books-are-written-on-mobile-phones> (accessed 15 December, 2007).
- Russell-Walling, Edward. "The Road to the Digital Future." *Publishers Weekly*, 12 March, 2007. URL: <http://www.publishersweekly.com/article/CA6423066.html?q=amazon+search+inside> (accessed 17 December, 2007).
- Sandals, Leah. "E-books Getting Closer: Booknet's Conference." *Quill & Quire Omni*, 23 March, 2007. URL: http://www.qullandquire.com/news/article.cfm?article_id=7665 (accessed 15 October, 2007).
- Sedgwick, Don. "Publishers Talk Digital Delivery at Google Conference." *Quill & Quire Omni*, 23 January, 2007. URL: http://www.qullandquire.com/news/article.cfm?article_id=7568 (accessed 6 December, 2007).

- Sellers, John. "Teen Marketing 2.0." *Publishers Weekly*, 3 September, 2007. URL: <http://www.publishersweekly.com/article/CA6473679.html?q=MYSpace&q=teen+marketing+2%2E0> (accessed 14 December, 2007).
- Shatzkin, Mike. "End of General Trade Publishing Houses." Presentation, Book Expo America. New York, NY, 31 May, 2007a. URL: <http://www.idealogs.com/endoftrade.htm> (accessed 14 June, 2007).
- . "Publishing and Digital Change: What's Next?" Presentation, Association of Book Publishers of British Columbia annual retreat, Qualicum Beach, BC, 11 February, 2006. URL: <http://www.idealogs.com/speeches/2006-02-11.html> (accessed 13 September, 2007).
- . "Success in a Parallel Universe: Perhaps With Some Help From Your DAD." Presentation, Book Industry Study Group's *Making Information Pay* conference, 10 May, 2007b. URL: <http://www.idealogs.com/speeches/mipdads.htm> (accessed 12 December, 2007).
- Skelton, Caroline. "Amazon.ca Debuts Search Tool, Minus Random House." *Quill & Quire Omni*, 12 August, 2005. URL: http://www.quillandquire.com/news/article.cfm?article_id=6819 (accessed 14 December, 2007).
- SlideShow Pro. URL: <http://slideshowpro.net> (accessed 25 November, 2007).
- Slowinski, F. Hills. *What Consumers Want in Digital Rights Management (DRM): Making Content as Widely Available as Possible in Ways that Satisfy Consumer Preferences*. New York City, NY and Washington, DC: Association of American Publishers, Inc. and American Library Association, 2003. URL: <http://www.publishers.org> (accessed 12 October, 2007).
- Stone, Brad. "Envisioning the Next Chapter for Electronic Books." *New York Times*, 6 September, 2007. URL: <http://www.nytimes.com> (accessed 17 September, 2007).
- Troy, Bethany. "Online Access to International Scholarly Content Set to Boost Canada's Research Capacity in Social Sciences and Humanities." *Press Release*. Ottawa, ON: Canadian Research Knowledge Network, 8 February, 2007. URL: <http://www.researchknowledge.ca> (accessed 12 July, 2007).

Vallik, Malle. "Publishers Off the Page: e-book/Digital Strategies." Presentation, BNC Technology Forum 2007. Toronto, ON, 21 March, 2007. URL: http://booknetcanada.com/events/bnc_tech_forum.html (accessed 12 October, 2007).

Visel, Dan. "Digital Editions." *Institute for the Future of the Book*, 20 June, 2007. URL: http://www.futureofthe-book.org/blog/archives/2007/06/digital_editions.html (accessed 6 October, 2007).

Wershler-Henry, Darren and Gregg Taylor. "Free-economics: Free Culture and the Future of Creative Content." Presentation, BookNet Canada Speaker Series. Toronto, ON, 13 September, 2007. URL: <http://www.booknetcanada.com/events/index.html> (accessed 4 October, 2007).

Wikipedia, "Web widget." URL: http://en.wikipedia.org/wiki/Web_widget (accessed 5 November, 2007).

William, Leigh Anne. "HarperCollins Canada adds soundtrack to Chandra novel." *Quill & Quire Omni*, 6 February, 2007. URL: http://www.quillandquire.com/news/article.cfm?article_id=7584 (accessed 13 December, 2007).

A P P E N D I C E S

APPENDIX A: E-BOOK FEATURES

Most frequently cited e-book features by users from the American Association of Publishers' DRM White Paper: *What Consumers Want in Digital Rights Management* published in 2003 (p.58).

Rank	Feature
1	Search
2	Personal copies (<i>Allow partial copying</i>)
3	Printing all or portion of work (<i>Allow partial printing</i>)
4	Institutional lending (<i>Lending</i>)
5	Interoperability (including <i>Format interoperability</i>) — Changing reading devices (<i>Portability among devices</i>)
6	Subscription based model — Encryption — Transmission/Downloading
7	Accessibility (<i>Access for blind and print disabled persons</i>) — Composite works — Personal lending (<i>Lending</i>) — Security — Site licensing
8	Consumer privacy — Consumer tools — Creating backups

APPENDIX B: CREATIVE COMMONS

From the Creative Commons website, the different types of licenses are the following:
(<http://creativecommons.ca/index.php?p=explained>)

Attribution:

You let others copy, distribute, display, and perform your copyrighted work — and derivative works based upon it — but only if they give you credit.

Noncommercial:

You let others copy, distribute, display, and perform your work — and derivative works based upon it — but for noncommercial purposes only.

No derivative works:

You let others copy, distribute, display, and perform only verbatim copies of your work, not derivative works based upon it.

Share alike:

You allow others to distribute derivative works only under a license identical to the license that governs your work.

These can be mixed-and-match to fit the different needs of various digital content.

APPENDIX C: DAM(D)S AND COVERSION PRICING

Gibson Library Connections Inc. provides in its document *Digital Asset Management and File Conversion for Canadian Publishers* an estimation for the pricing on ongoing usage of DAMDs (available at <http://www.gibsonlibraryconnections.ca/codeMantraFAQ.pdf>):

DAM(D)s

Prices are based on the number of titles stored and managed by the system. The minimum annual fee is \$2,600, which covers system use and storage of 100 titles. Should the licensed quantity be exceeded during a year, a new invoice will be generated for an additional quantity at the applicable discounted price. (That is, if a client were to enter an initial quantity of 100 titles and then require a further increment of 100, the applicable invoice amount would be \$24.50 per title.)

Titles loaded	System Access		Archiving		Total
	Per title	Subtotal	Per title	Subtotal	
Up to 100	15.00	\$1,500.00	11.00	\$1,100.00	\$2,600.00
101 to 200	14.00	\$2,800.00	10.50	\$2,100.00	\$4,900.00
201 to 300	13.00	\$3,900.00	10.00	\$3,000.00	\$6,900.00
301 to 500	12.00	\$6,000.00	9.50	\$4,750.00	\$10,750.00
501 to 800	11.00	\$8,800.00	9.00	\$7,200.00	\$16,000.00
Over 800	10.00	\$10,000.00	8.50	\$8,500.00	\$18,500.00

Universal PDF¹ File Conversion

Output format	Price per title
Universal PDF	\$90.00
Universal PDF and codeMantra Handheld Suite (MobiPocket, MS Reader/LIT and Palm)	\$120.00
Universal PDF and codeMantra Handheld Suite and print-on-demand	\$150.00
codeMantra Handheld Suite only	\$90.00
Print-on-demand only	\$90.00

¹ Universal PDF is codeMantra's version of PDF conversion that includes all the unifying elements and needs of the different e-book distributors.