

**THE TANGLED WEB: MANAGING AND
CONFRONTING SCHOLARLY EBOOK
PRODUCTION AT UBC PRESS**

by

Alison Elaine Knight
M.A, University of Victoria, 2006
B.A. (Hons), University of British Columbia, 2004

PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF PUBLISHING

In the
Faculty of
Arts & Social Sciences

© Alison Elaine Knight 2007
SIMON FRASER UNIVERSITY
Summer 2007

All rights reserved. This work may not be
reproduced in whole or in part, by photocopy
or other means, without permission of the author.

APPROVAL

Name: Alison Elaine Knight
Degree: Master of Publishing
Title of Project: The Tangled Web: Managing and Confronting
Scholarly Ebook Production at UBC Press

Examining Committee:

John Maxwell
Senior Supervisor
Assistant Professor, Master of Publishing Program

Rowland Lorimer
Supervisor
Director, Master of Publishing Program

Holly Keller
Industry Supervisor
Assistant Director
Editorial and Production Services, UBC Press

Date Approved:

August 14, 2007



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

ABSTRACT

This report describes academic ebook production as it currently stands, emphasizing the practices of UBC Press and its use of a digital asset management/distribution provider. It identifies best ebook production practices, emphasizing labour and cost reduction, and argues for the replacement of digital asset management/distribution services with industry standardization.

Part One identifies the organizations and technologies involved in scholarly ebook production, covering formats, readers, access, digital rights management, and distributors. It describes the ways in which these areas intersect to necessitate short-term digital asset management and long-term standardization.

Part Two examines UBC Press ebook production, outlining its adoption of codeMantra's digital asset management services. It analyzes codeMantra's facilitation of the creation, distribution, and management of UBC Press ebooks in the face of convoluted industry requirements.

Part Three identifies potential enhancement to ebook production, focusing on standardization initiatives that could obviate the need for digital asset management.

Keywords: Ebooks; Ebook production; Digital asset management/distribution; Standardization; UBC Press

Subject Terms: Electronic Books; Electronic Publishing; Scholarly Publishing; Standardization

ACKNOWLEDGEMENTS

Thanks are due to the staff members of UBC Press, who were extremely supportive during the formation of this report; Holly Keller, Darcy Cullen, Anna Eberhard Friedlander, and Peter Milroy deserve particular acknowledgement. Thanks also to Bob Gibson, of Gibson Publishing Connections. I would also like to extend my gratitude to the Canadian Centre for Studies in Publishing, most particularly John Maxwell, Rowland Lorimer, and Jo-Anne Ray, for their considerable efforts towards the improvement and completion of this report.

TABLE OF CONTENTS

Approval.....	ii
Abstract.....	iii
Acknowledgements	v
Table of Contents	vi
List of Abbreviations	viii
Introduction.....	1
1: Ebook Production: An Overview	5
1.1 Formats.....	6
1.1.1 Markup Language	7
1.1.2 PDF.....	9
1.2 Reading Devices.....	10
1.2.1 Web Browsers.....	10
1.2.2 Adobe Reader	11
1.2.3 Proprietary Readers.....	11
1.3 Methods of Access	12
1.3.1 Simple Web Access	12
1.3.2 Tethered Access.....	13
1.3.3 Downloads	14
1.4 Digital Rights Management	14
1.5 Ebook Distributors.....	16
1.5.1 NetLibrary.....	19
1.5.1.1 Methods of Access	19
1.5.1.2 Rights to Access.....	20
1.5.1.3 Submission Requirements	22
1.5.2 Ebrary	23
1.5.2.1 Methods of Access	24
1.5.2.2 Rights to Access.....	24
1.5.2.3 Submission Requirements	26

1.5.2.4 Canadian Electronic Library	26
1.6 Ebook Production Summary: Digital Asset Management/ Distribution and Standardization.....	27
2: Ebook Production at UBC Press	31
2.1 History and Current Status of UBC Press Ebooks	31
2.2 UBC Press and Digital Asset Management/Distribution.....	34
2.2.1 Ebook Creation Services	35
2.2.1.1 Universal PDF	35
2.2.1.2 Print on Demand	37
2.2.2 Collection Point	38
2.2.2.1 File Identification	40
2.2.2.2 File Transfer	45
2.2.2.3 File Tracking	49
2.2.2.4 Archiving	52
2.2.2.5 Collection Point Summary and Suggestions.....	53
2.3 Financial Summary of Ebook Production at UBC Press	53
2.3.1 Ebook Creation.....	54
2.3.2 Collection Point.....	55
2.3.3 Revenue.....	55
2.4 Ebook Production at UBC Press: Conclusions	56
3: Standardization Initiatives.....	58
3.1 International Digital Publishing Forum: Open Publication Specification Standard.....	58
3.2 Open Standard PDF	60
3.3 Digital Rights Management Initiatives.....	61
3.4 Association of Canadian Publishers	64
3.5 Canadian Research Knowledge Network’s Digital Content Initiative.....	65
3.6 Standardization Conclusions	67
4: Conclusions	69
Appendices	72
Appendix A: Ebook Format Types.....	73
Appendix B: Ebook Distributors	74
Appendix C: Freelancer Upload Guidelines.....	75
Reference List.....	78

LIST OF ABBREVIATIONS

ML	Markup Language
PDF	Portable Document Format
uPDF	Universal Portable Document Format
HTTP	HyperText Transfer Protocol
FTP	File Transfer Protocol
DAM	Digital Asset Management
DAD	Digital Asset Distributor
DAMD	Digital Asset Manager and Distributor
DRM	Digital Rights Management
POD	Print on Demand

INTRODUCTION

In the late 1990s, the ebook was widely expected to revolutionize publishing. With its promise of immediate availability, low costs, and state-of-the-art features, industry analysts like Andersen Consulting predicted the development of a multibillion dollar ebook market by 2005.¹ Yet the reality for most publishers of ebooks has been one of confused production and low revenues. The reasons behind this state of the ebook market are manifold. Underdeveloped e-reader technology, lack of reader willingness to abandon print, and widely divergent file format requirements on the part of distributors and readers combine to produce a market defined by inflated disorder and wary indifference.

Yet, despite lacklustre beginnings, the ebook is becoming an important resource for publishers. This is particularly true in scholarly publishing, for which market conditions are quite different than for trade publishing; University presses are traditionally small-scale operations serving focused markets powered

¹ Andersen Consulting (now Accenture), "Reading in the New Millennium: A Bright Future for E-book Publishing" (New York: American Association of Publishers, 2000); summary at http://www.accenture.com//rdonlyres/78BBA6F9-7370-49A3-AEFF-46C77CC6669A/0/medi_aap.PDF (accessed 29 July 2007).

by libraries, rather than diverse markets of bookstores and individual purchasers. While ebooks do not appear any more likely to eclipse print in scholarly publishing than in trade, the sizeable purchasing power of libraries looking to digitize combined with the benefits of e-distribution, decreased production costs, and technological innovation offer enough of a measurable benefit to make ebooks a viable addition to print activities. In response to this potential for scholarly ebooks, many scholarly publishers are eager to develop new ebook strategies.

Scholarly publishers must adopt ebook strategies that address the problem of widely divergent file format requirements on the part of both consumers and distributors. In order to provide ebook offerings through all available channels and to all consumers, publishers must produce ebook files in multiple formats, which usually involves costs and labour that currently outweigh the revenue that ebooks are capable of producing. Increasingly, publishers realize that it is through streamlining the ebook production process that they can best serve the needs of consumers and distributors, as well as make ebooks a valuable resource for their organizations. For many publishers, the best method of streamlining ebook production is to utilize a Digital Asset Management and Distribution (DAMD) service, which organizes the multiple files and formats required by the many parties involved in the ebook market.

This report outlines the ebook production process as it currently stands in academic publishing, focusing on the technologies and organizations that shape its current complexity. In particular, it describes the developments in ebook production made at UBC Press, arguing for the usefulness of its new ebook creation and DAMD service provider, codeMantra. At the same time, it argues that while DAMD services improve scholarly ebook production, they are not a long-term solution to the convulsion of the current ebook market. A long-term solution must instead be sought through the establishment of industry standards.

Part One describes the multiple formats, reading devices, methods of access, Digital Rights Management (DRM) technologies, and distributor requirements that combine to produce the ebook market's current complexity. Throughout this description, the importance of open and interoperable standards is highlighted.

Part Two moves to a study of UBC Press's adoption of codeMantra's DAMD service, and its improvement of the Press's ebook production process. This study suggests that UBC Press's decision to use a DAMD is an effective means of addressing the ebook market's complexity, but also that the use of a DAMD service is an implicit reinforcement of that complexity.

Part Three suggests that a long-term solution to the complex requirements of the ebook market must come from the establishment of open format industry standards, primarily through the actions of standardization initiatives and cooperative industry programs.

1. EBOOK PRODUCTION: AN OVERVIEW

All books begin as ebooks. Books are no longer composited and printed using physical blocks of type; rather, they are transferred to printers in the form of digital files. Publishers rely on digital formats to transmit and manipulate works throughout the publishing process, to the extent that many books are not embodied on paper until they have been through the printer's shop. In this sense, the book production process and the ebook production process overlap, and the creation of an ebook cannot be separated from the production of the book as a whole.

Ebook production becomes an issue of translating a book's textual content into digital formats that can be easily distributed to e-distributors and readers, and easily managed by press staff. Decisions made about the creation of an ebook underwrite its capabilities for distribution and use, and are underwritten by the ways in which publishers manage the digital content that they wish to disseminate. In this way, the creation, distribution, and management of ebooks become inseparable components within a total process.

In order to ensure that informed decisions are made in all of these three areas, it is essential for ebook producers to understand the technologies and organizations involved in generating, distributing, and organizing their ebook content. The following describes the file formats, reading devices, and methods of accessing, protecting, and distributing ebooks available to ebook publishers, as well as the impact that their current status has on ebook production.

1.1 Formats

Format specifics such as size, cover, paper type, or font can vary almost infinitely in traditional print production, but these variations will have relatively little impact on a book's capabilities for use; readers are equally able to read from a 5"x8" paperback as they are a 6"x9" hardcover. In contrast, variations of format have enormous repercussions for an ebook's usability. Ebook production is complicated by the multiple file formats available; texts can be digitally embodied within near-infinite variations of format, variations that carry significant implications for use (for a full list of ebook format types, see Appendix A).

Although specifics differ widely, in the current market, ebooks are generally offered in either Markup Language (ML) format (that is, HyperText Markup Language [HTML], eXtensible Markup Language [XML], and eXtensible

HyperText Markup Language [XHTML]) or PDF format. With the addition of programs such as Java or Flash to MLs and interlinking to PDFs, as well as variations between the specifics of files (such as size or security protection), the ebook market's preference for these two format types does little to simplify matters. However, it does provide ebook producers with a point of entry.

1.1.1 Markup Language

Much of traditional Web publishing is encoded in Markup Language (ML), which forms the bulk of website content. ML's potential as a medium for scholarly publications has been tested by many digital projects; for example, the University of Victoria's Internet Shakespeare Editions website (ise.uvic.ca) offers XHTML-based editions of Shakespeare's works and associated criticism, and PubMed (www.pubmed.gov) offers XML-based articles addressing new research in the medical field.

ML is appealing to publishers because of its flexibility. ML is in essence *information* about a given text; an ML file provides both the text and directions about what to do with it. All that a publisher need do is change those directions (for example, concerning font size and type) and the text is correspondingly changed. The fact that ML-based ebooks are defined by flexible rules about the text they contain also means they are "reflowable"; that is, they do not have

predetermined line or page breaks, so like water into jars they conform to different screen readers. With the inclusion of interactive programming, ML books can permit scholarly publishers to use that flexibility to move away from traditional page and codex formats, in a manner that re-evaluates the book as a medium. ML-based books do not need to be read in a linear fashion; rather, they can be made fully linked and searchable, available to be read in a connection with other works and other parts of themselves. Lastly, ML-based works are flexible regarding the complexity of the works that they support; works can be made very complex and highly functional, or they can be made simple, allowing viewing through handheld devices.

ML-based ebooks are not without failings; from a production standpoint MLs frequently present problems. Creating ML documents requires individuals highly trained in encoding, a requirement that often exceeds publishers' IT capacities. Also, for many publishers, the flexibility of ML display is in conflict with the importance of maintaining control over text design.² Divergence in end-user systems frequently results in irregularities of screen display, from lineation and pagination to typeface and the placement of images. Publishers who are

² The control required by publishers over text design speaks to the importance of good typography in the creation of a book. The influence that text layout and visual appearance can have over the reader's experience is considerable, and therefore must be carefully tuned by publishers.

dedicated to design standards and insist on controlling the presentation of their ebooks often turn instead to the Portable Document Format, or PDF.

1.1.2 PDF

For most scholarly publishers, the ebook gold standard is PDF. Whereas MLs establish flexible rules for their texts, rules that are then translated into display, a PDF document establishes the specific layout of text upon a page. Developed by Adobe, PDF is a page-representation format that displays text and images as they have been laid out in an original document, like a virtual snapshot. Where MLs allow flexibility, PDFs ensure control. PDFs are also defined by their ease of use; unlike books published in ML, the creation of which requires considerable technical expertise, PDFs can be easily generated by publisher staff or freelance designers.³ As the format that is most frequently sent to printers, PDF ensures a press-quality display without requiring additional processing. PDF is not without flexibility: it can support inter- and intralinking, 3D graphics, audio, and end-user highlighting and commenting features. Indeed,

³ Although levels of technological expertise vary considerably between scholarly press employees and freelancers, at houses such as UBC Press, where all regular employees are able to use programs with difficulty levels similar to Microsoft Office (if not significantly higher), the creation of a PDF presents no problem. Furthermore, any typesetter (who is usually responsible for the creation of PDFs in most production departments) will be proficient with the Adobe platform.

PDF unites a degree of the flexibility of ML with design control and ease of use. As such, it is the preferred format for most ebook producers.

1.2 Reading Devices

Any type of hardware or software intended to display ebook content is a *reading device*. As a general means of classification, reading devices can be categorized as either PC-based or portable; in many instances, however, reading devices are rarely restricted to a single scenario of use. In general, portable devices are targeted for trade publications, if they are used at all. Whether portable reader technology or reader tastes are at fault, portable readers have been decidedly ignored by consumers. Scholarly ebooks are read almost exclusively through PC-based Web browsers and Adobe Reader, although some must be read through proprietary reading devices.

1.2.1 Web Browsers

Familiar to almost all computer users, Web browsers display content encoded in MLs via the Internet. A browser translates the markup, or directions about a text, and then displays only the text, per those directions. ML-based content is not Web browser specific; an XML file, for example, is equally readable

through both the Mozilla Firefox and Internet Explorer browsers. ML ebooks will, however, vary in their expression across Web browsers.

1.2.2 Adobe Reader

The Adobe Reader is the universal reading device used to view PDF files. Unlike ML content, which can be viewed through many different browsers, PDFs can only be read in Adobe Reader (or Adobe Acrobat, which allows PDF creation and manipulation). Fortunately, the blanket requirement for the Adobe Reader when viewing PDFs is not complicated by undue barriers to access: Adobe Reader is almost always included with new computers, and if it is not, Adobe Reader is free to download from the Adobe website (<http://www.adobe.com>). Although many of the features of PDFs made possible by the most recent version (Adobe Reader 8) are not supported by earlier versions, the ubiquity of Adobe Reader still ensures a more regular expression of an ebook's features.

1.2.3 Proprietary Readers

Some ebook distributors and technology providers offer proprietary reading software for their ebooks, which are intended to be used in place of Adobe Reader or a Web browser. The most notable example of a proprietary reader in scholarly ebook publishing is the ebrary Reader, which is required to

read ebrary's PDF-derived Exchange Document Format (EDF) ebooks.⁴ Another is the MobiPocket Reader, which supports ebook viewing on most major handheld Personal Data Assistant (PDA) devices (as PDAs often do not have the memory capacity to display a full Web browser or Adobe Reader). Both of these reading devices are free to download, so they are not difficult for users to access; however, as proprietary readers are partnered with proprietary and closed formats (formats that do not work across multiple platforms), they contribute to the multiple format needs of the ebook market.

1.3 Methods of Access

Ebook content, whether it is in ML or PDF format, is typically transmitted to users over the Internet.⁵ Both are amenable to many different browser types and configurations. There are, however, key differences between ML and PDF methods of access. These are described in the following sections.

1.3.1 Simple Web Access

Web browsers display ebook content by communicating with a host computer that contains the ebook's files. Using HyperText Transfer Protocol

⁴ Publishers need not concern themselves with producing EDF documents, as this activity is undertaken by ebrary; see section 1.5.2.3 herein for more information.

⁵ Occasionally, ebook providers will send ebooks to users as CD-ROMs, but this is quite uncommon.

(HTTP), an end-user's computer (called a *client*) asks the host to transmit ML content. Once ML content has been transmitted, end-users are able to view text and graphics (and, if they choose, markup⁶). ML files are accessible only when the user is connected to the Internet, unless the user saves the ML content (page by page) within their computer for offline viewing.

1.3.2 Tethered Access

With "tethered access," ebook content is transmitted to users page by page, via the Internet. Ebook providers using tethered access create a separate webpage for each page of an ebook; as a result, users access a series of fixed (not reflowable or reformattable, similar to PDF) page views that represent the way the book would look in print. This is done so that users will not have access to an ebook's source files, but instead are only able to view representations of individual pages (imagine viewing a page through a glass window); it also ensures that readers do not have access to the file itself, which could enable copying. Tethered access files cannot be saved, copied, or manipulated; however, providers often allow users to print and/or copy and paste a portion of that file's text, in accordance with copyright law.

⁶ This is achieved by viewing the page source.

1.3.3 Downloads

Alternatively, users can access ebook PDFs by downloading whole ebook files. When ebooks are accessed in this manner, copies of the files themselves are saved in the end-user's computer, enabling users to read, modify, or comment on them. If a tethered access PDF is similar to single pages seen through a window, then a downloaded PDF is like a copy of the entire book sent in the mail; tethered access content is no longer available after the connection is severed, but downloaded content remains within the end-user's computer. The control that downloading allows the end-user generally requires ebook providers to include some type of copy protection.

1.4 Digital Rights Management

Tethered access and copy protection represent two strategies for protecting ebooks from unauthorized use, otherwise known as Digital Rights Management (DRM). DRM is intended to protect slim profit margins (particularly important for scholarly publishing), as well as to "create an essential foundation of trust between authors and consumers."⁷ By providing

⁷ See American Association of Publishers, *Digital Rights Management for Ebooks: Publisher Requirements* (New York: American Association of Publishers, 2000), <http://publishers.org/digital/drm.PDF> (accessed July 31, 2007), 6; also cited in Stephen Mooney, "Interoperability: Digital Rights Management and the Emerging Ebook Environment," *D-Lib Magazine* 7. 1 (Jan 2001), <http://www.dlib.org/dlib/january01/mooney/01mooney.html> (accessed July 31, 2007).

works through technologies that prevent users from illegally accessing, copying, manipulating, or transferring files, ebook providers (publishers, distributors, libraries, and authors) protect their investments of time and resources from exploitation.

In order to restrict unauthorized access to ebooks, providers utilize a variety of techniques; most commonly, DRM takes the form of ebook encryption, user authentication (password protection and digital certification), and secured ebook storage and transmission. Commonly, ebook producers integrate these methods of DRM through proprietary technology and closed formats. It is this practice that makes DRM a double-edged sword, as proprietary (i.e., non-standard, unique to each provider) formats render ebooks incompatible across platforms, inaccessible to many would-be readers, and not durable as platforms change, contributing to the confused and multiformatted status of the ebook market.

Most ebook producers, and especially scholarly publishers, are committed to producing works that are enduring and accessible, and hope to do so in balance with protecting fragile revenues. For this reason, the ebook industry is gradually incorporating DRM with open formats (formats that work across

platforms and can be manipulated) and even establishing new forms of DRM⁸ (for more on this industry push, see section 3.3). Ebook providers appear to appreciate the fact that factious technologies do not protect profits, but rather drain resources and prevent the widespread adoption of ebooks.

1.5 Ebook Distributors

One of the primary advantages of ebooks for scholarly publishers is the ease of their distribution: publishers need not coordinate with printers, wholesalers, delivery services, bookstores, and libraries in order to make sure that end-users receive books. Rather, ebook files are transferred from publishers to customers through e-distributors, who assume many of the above roles.

The fundamental concerns of ebook distributors are not only selling, but also providing and policing *access* to ebook content. E-distributors are responsible for making ebook content available to authorized customers and ensuring that ebooks are protected from unauthorized use (i.e., use without purchase). In their attempts to best provide and police access, ebook distributors differ widely in the methods of and rights to access that they offer. Some operate like traditional booksellers, allowing customers to purchase individual books. Others have more in common with purchasing consortia — they develop a

⁸ See American Association of Publishers, *Digital Rights Management for Ebooks*.

collection of works and offer it to customers as a whole. Still others behave like Publisher's Clearing House agents, offering temporary subscriptions to books; finally, some operate like amusement park ticket agents, allowing unlimited access to any title in their collection for a fixed length of time.

Publishers currently have over twenty ebook distribution services to choose from (see Appendix B for a complete list). The sale of scholarly ebooks is primarily directed at libraries rather than single end-users. The library operates as the e-distributor's customer, purchasing rights to access works, which are then made available to patrons. As a result, the major organizations involved in scholarly e- distribution are library aggregators. These distributors provide as part of their service the aggregation of content suitable for libraries.⁹ Focusing on libraries presents another complication for providing and policing access, as the access granted to individual library patrons must also be coordinated. In order to gain a representative view of library ebook distributor types and an understanding of the impact that meeting their needs can have on the ebook production process, the following section explores two major library aggregators, NetLibrary and ebrary. Scholarly presses certainly offer e-content through a

⁹ All subsequent references to library aggregators and aggregation will therefore refer jointly to the distribution (selling, providing, and policing access) and aggregation functions performed by library aggregators.

variety of other channels (including other e-distributors, such as Baker & Taylor or MyiLibrary, as well as ebook preview services such as Google Book Search and Amazon's Look Inside), but these two organizations present the common characteristics of e-distributors in the scholarly market and their requirements.

Before beginning a discussion of NetLibrary and ebrary, it is useful to define the industry terms used to describe common access models:

Perpetual Access: The perpetual right, granted to an authorized customer (e.g., a library), to make a title available to authorized end-users (e.g., library patrons). Perpetual access rights are similar to the rights obtained by purchasing a book at a bookstore. Perpetual access can be granted title by title, or customers can acquire perpetual rights to a library aggregator's entire collection.

Subscription Access: The limited-term right, granted to an authorized customer, to make ebooks available to end-users. Subscriptions can also be purchased on a title-by-title basis or for entire aggregator collections. Subscription access is not like a magazine subscription, in which a new issue arrives at fixed intervals; rather, subscribers to ebooks can access those ebooks as many times as they like until the subscription term ends.

Single Viewing: Allows only one authorized end-user (patron) to gain reading access to a title at a time. As with traditional print library collections, while another patron is using a work, it is not available to others. This is the usage model preferred by most scholarly presses, as it protects course print sales.¹⁰

Multiple Viewing: Allows multiple authorized end-users to gain reading access to a title at a time.

1.5.1 NetLibrary

NetLibrary was one of the first organizations to offer aggregation services to publishers; it has since grown to become the leading ebook library aggregator.¹¹

1.5.1.1 Methods of Access

NetLibrary does not provide ebook content to library customers directly, rather, after a library has purchased rights to a title, that library's *patrons* are able to access titles through the library's dedicated NetLibrary site. Patrons select

¹⁰ By making a copy available to only one reader at a time, publishers ensure that all students within a course will not depend on the library copy for books set as required reading.

¹¹ All information about NetLibrary has been drawn from <http://www.netlibrary.com>, accessed both externally and through UBC Press's dedicated NetLibrary website. All information included in this report is accessible publicly; no confidential material has been imparted in this study.

ebooks from their library's collection on this NetLibrary site, ebooks that they can then read via tethered access in XHTML format, using JavaScript and frames. Some libraries also purchase an Adobe Content Server Gateway License, which allows authenticated users to download PDF versions of books to read offline.

1.5.1.2 Rights to Access

NetLibrary offers two models for access rights: title-by-title perpetual access, or subscriptions to entire collections. Either can be purchased by a library at a cost of list price plus 55% for perpetual access or 15% for a subscription (a percentage of which is remitted to publishers as royalties¹²). If a library has purchased rights to an ebook, its authorized users are permitted to access that ebook. Individuals become authorized users by creating a library-affiliated NetLibrary account, which can only be done through the library's dedicated NetLibrary site (in general, these networks can be accessed only by library patrons). Authorized users must then log in to this account before they can access ebook content.

¹² Because royalty structures are negotiated with each publisher in the form of confidential agreements, they are not available for this study.

NetLibrary offers a single viewing access model under which users are permitted fifteen minutes of active viewing per session. Following this fifteen minute session, or if the user ceases active reading (i.e., turning pages), the title becomes available for other users (the initial user is permitted back-to-back sessions if others have not accessed the title). Users are permitted to copy and paste and/or print sections of text, per copyright law (in Canada, up to 10% or one chapter of an entire work, whichever is greater¹³). Unlike print books, for which copyright violations are difficult to police, the NetLibrary platform enforces copyright law, as its DRM technology prevents users from exceeding copy and print limits per session.

Libraries that have acquired an Adobe Content Server Gateway License are permitted to make titles available for offline reading. Authenticated users of libraries with this license are able to store a copy of the ebook PDF file on their own computer for a fixed term (determined by the library). Much as with a software trial, this PDF file cannot be opened after the “loan” period has passed. This PDF copy is available exclusively for viewing (its DRM protection prevents copying and printing), but users can elect to view it on their home computers or transfer it to a handheld reading device. Users must have Adobe Reader version

¹³ Access Copyright, “Incidental Photocopying,” *Access Copyright*, <http://www.accesscopyright.ca/Default.aspx?id=93> (accessed July 20, 2007).

6.0 or higher in order to view NetLibrary PDFs; however, as Adobe Reader is available online for free downloads (currently offering version 8), this is generally not a problem for users.

1.5.1.3 Submission Requirements

Publishers who distribute through NetLibrary must provide files that satisfy the following requirements, outlined in its submission guidelines¹⁴:

<i>PDF:</i>	Version 1.3 or higher
<i>Software:</i>	Adobe Acrobat 4.0 or higher
<i>Text:</i>	Single document with entire book contents (front matter, sections, parts, chapters, end matter, blank pages)
<i>Cover:</i>	Front cover only 400 pixels wide 72 dpi minimum
<i>Images:</i>	Colour, Grayscale, Bitmap: 72 dpi minimum
<i>Fonts:</i>	Seven bit ASCII (Western European). Embedded
<i>TOC:</i>	Bookmarked Standard fonts Western European encoding Linked to contents
<i>Pagination:</i>	Uniquely labelled pages ("cover" and numbers)
<i>Disclaimers:</i>	Omitted text and images
<i>Cropping:</i>	Uniform Minimal "dead space" No registration or printer marks
<i>Orientation:</i>	No landscape pages - rotate content to vertical
<i>Encryption:</i>	Remove Adobe DRM

¹⁴ NetLibrary submission requirements compiled by author from information contained in NetLibrary, *PDF Submission Guidelines*, <http://a1835.g.akamai.net/7/1835/276/95d75154a85342/extranet.netlibrary.com/extranet/publisher/specifications/documents/PDFSubGuidelines.PDF> (accessed July 18, 2007).

<i>Metadata:</i>	Publisher	Title
	Editor/Author(s)	Print ISBN
	Publisher Ebook ISBN	Publication Year
	Price	File Format
	Bookmarks	Page Labels
	Image Electronic Rights Y/N	Images to Exclude (Page #)
	Cover Detached Y/N	World Electronic Rights Y/N
	Country Exceptions	

Submissions that fail to meet these requirements are converted by NetLibrary at a fixed per-book charge of \$110 USD. The extent of the requirements of this single distribution platform (e.g., cover dpi, cropping, linking) illustrates the magnitude of the labour involved in tailoring ebook files to multiple distributors.

1.5.2 Ebrary

Ebrary is a major figure in ebook distribution, but not only because of its role as a library aggregator.¹⁵ Rather, ebrary licenses not only access to its ebook collection but also the use of its platform. The platform is often used by other ebook distributors; more importantly, the platform is adopted by ebrary library customers, which allows them to augment their electronic collections as they desire (for example, university libraries might offer patrons electronic versions of students' dissertations or purchase directly from publishers the rights to ebooks

¹⁵ All information regarding ebrary has been gathered from <http://www.ebrary.com>, accessed both externally and through the Simon Fraser University library online catalogue at <http://www.lib.sfu.ca>. All information included in this report is accessible publicly; no confidential material has been imparted in this study.

not offered by ebrary). Recently, several large publishers (Elsevier, Kluwer, Oxford University Press, Gale, and Wiley) acquired the ebrary platform for use in distributing their own ebooks, thus becoming both ebook publishers and ebook distributors.

1.5.2.1 Methods of Access

Because libraries acquire the ebook platform, users do not have to search for books through a separate website (as with NetLibrary); instead, users search for books within the library's own catalogue site, and works available as ebooks will have a link within the catalogue record. Once a user has selected an ebook to view, the content is provided through tethered access in ebrary's Exchange Document Format (EDF), which provides each page with its own URL for tethered access viewing using the ebrary Reader. The ebrary Reader is free to download, and it maintains many of the features of Adobe Reader (such as commenting and highlighting), with the addition of personal bookshelf, multiple field search, and automatic citation features.

1.5.2.2 Rights to Access

Ebrary's licensing is based on flexible models: publishers can elect to provide subscription, perpetual access, or combined purchase rights, as well as flexible viewer usage models. Similarly, libraries can choose which specific types

of rights they wish to purchase (prices are dependent on the library's *full time equivalent*, which is essentially its total number of authorized users). For subscriptions, ebrary offers several packages, from its Academic Complete collection to single subject subcollections (or subcollection combinations). Ebrary also allows libraries to subscribe solely to the lists of specific publishers. The flexibility of this model has made ebrary's platform increasingly popular; rather than impose blanket rights models, ebrary allows publishers and libraries to determine which rights they would like to release or acquire on a title-by-title basis if desired.

End-user rights to ebrary content are, like NetLibrary's, dependent on an account at the host library. End-users gain access to ebrary content through password authorization; with many university libraries, an account is created when the patron establishes an institutional computing ID. If authorized by the library to view content, patrons may do so one page at a time. Because ebrary ebooks are viewed page by page via tethered access, readers never have access to the ebook files themselves. This restriction enforces copyright law, as ebrary Reader DRM permits users to copy and paste or print only a maximum of 5% of the total document (the equivalent of a few pages).

1.5.2.3 Submission Requirements

In order to produce the ebrary EDF format, publishers must provide ebrary with files meeting quite specific guidelines in terms of file optimization, security settings, fonts, images, page size, crop marks, and accompanying metadata. Ebrary supplies publisher partners with these guidelines in a twenty-six page brochure, which is confidential and thus not available for inclusion in this report.

1.5.2.4 Canadian Electronic Library

As outlined above, ebrary is not only a distribution service. It also licenses its platform to power other ebook distribution or collection services, such as the Canadian Electronic Library (CEL). CEL is an initiative spearheaded by Gibson Library Connections and Gibson Publishing Connections, the Canadian operator of codeMantra's Collection Point (a Digital Asset Management and Distribution [DAMD] service to be explored in detail in Part Two of this report). CEL aims to develop the premiere collection of Canadian scholarly publications, aggregated for distribution to Canadian libraries. This service offers a joint value to libraries and scholarly publishers, as the specific focus on Canadian scholarly monographs and collections allows them to operate like a purchasing

consortium.¹⁶ Because it uses the ebrary platform, CEL's methods and rights to access follow the ebrary model.

1.6 Ebook Production Summary: Digital Asset Management/Distribution and Standardization

The variation in formats, reading devices, means of access, and distributor requirements demonstrates the enormous range of ebook options that producers must coordinate. Variation in distributor requirements especially, such as the many slight differences between ebrary submission guidelines and those of NetLibrary (e.g., pixel width requirements for covers, etc.), illustrate the extent to which ebook production impacts a wide range of activities. Press production departments cannot produce a single ebook file suitable for all parties involved, as with print; rather, the needs of distributors, libraries, and end-users must be written into the document files themselves, many times over as they are multiplied across a publisher's entire list.

Similarly, the ebook files produced for these multiple players are by no means single documents; rather, they grow out of many component files as the ebook moves through manuscript editing, typesetting, proofreading, cover

¹⁶ See Gibson Library Connections, "Canadian Electronic Library," *Gibson Library Connections*, <http://www.gibsonlibraryconnections.ca/CPC%20Partners%20FAQ%20Sheet.htm> (accessed July 13, 2007).

design, and more. With this accumulation of key files, ebook production comes to mean many things, not the least of which is management. Management involves controlling what files are in use, who they are given to, and how they are manipulated, ensuring that all parties (including the production department itself) receive what they need.

Many larger university presses (such as Cambridge or Oxford) have IT departments dedicated to coordinating their operational data; others (such as UBC Press) depend on select individuals and specific IT software (Klopotek, Acumen). Even so, coordinating the *data* involved in a publisher's operations does not necessarily include managing *digital assets* (i.e., the content files comprising a finished title) as part of the production process. It is for this reason that ebooks (and, indeed, print books) require a production-specific Digital Asset Management (DAM) system. For many publishers, this system can be efficiently developed in-house. Many others, however, are beginning to turn to outsourced DAMs for streamlining the management of digital assets. Many such DAMs also involve Digital Asset Distribution systems, or DADs; DADs link with all of the parties involved in the production and use of a digital asset, ensuring that all

parties receive the files they need, when they need them and in the proper formats.¹⁷

An analysis of UBC Press's transition to an external Digital Asset Management and Distribution (DAMD) service will be the focus of Part Two of this study. For many publishers, UBC Press included, external DAMD service providers offer a necessary function; not only do they help coordinate the use of files and the communication between parties involved in ebook creation, management, and distribution, they do so with far more digital expertise and manpower than a typical scholarly press possesses. As the following section will demonstrate, external DAMDs are a necessary and effective means of addressing the confusion and multiplicity of the current ebook market, in a way that is far superior to previous practices.

It should be noted, however, that DAMDs are not a final solution. Rather, they encourage e-distributors to maintain the convolution of current arrangements. Before the ebook market establishes concrete roles and rules for its many players, it will become necessary to work towards eliminating the need for DAMDs altogether by developing alternative solutions, such as industry

¹⁷ It is important to make the distinction between e-distributors and digital asset distributors (DADs); DADs do not perform the distribution functions outlined in section 1.5; rather, they facilitate the transfer of files from publishers to others (including e-distributors).

standards for formats, creative DRM, and cohesive distribution practices.

Initiatives intended to provide long-term ebook solutions will be examined in

Part Three.

2. EBOOK PRODUCTION AT UBC PRESS

UBC Press's experience with ebooks parallels that of many scholarly publishers, in that they are committed to upholding scholarly and production standards and eager to support technological innovation and the responsible dissemination of digital content, but limited to doing so in a way that does not damage print book sales or incur undue costs. As such, the success of UBC Press's ebook activities hinges largely on the efficiency of the production department. UBC Press has recently attempted to enhance ebook production through the integration of an external DAMD service, provided by codeMantra. The Press's experience demonstrates the significant simplification of the ebook production process enabled by DAMD services, suggesting the effectiveness of DAMDs as a short-term means of addressing the confusion of the ebook market.

2.1 History and Current Status of UBC Press Ebooks

UBC Press began producing ebooks in the late 1990s in partnership with NetLibrary. At this time, PDFs were not yet a standard feature of the print-book production process, so ebooks had to be created by scanning print copies. UBC Press agreed to allow NetLibrary to perform these scans and distribute the

finished ebooks (at this time, in HTML format) in exchange for a percentage of royalties. However, in NetLibrary's original iteration, UBC Press found that the HTML format resulted in frequent pagination problems, requiring Press staff to expend significant labour vetting finished ebooks. Furthermore, UBC Press was dissatisfied with NetLibrary's royalty structure (which at the time disregarded UBC Press's pricing instructions, and offered all ebooks at paperback prices; deducted 50% of profits, without – in the opinion of UBC Press managers – adding commensurate value; and failed to fulfill delivery of complete royalty payments¹⁸). After 2003, UBC Press elected to terminate its partnership with NetLibrary. With the general adoption of the PDF throughout the publishing industry, UBC Press moved to ebrary and Baker & Taylor (the world's largest library distributor, which offers HTML-based books through dedicated library sites much like NetLibrary's) for ebook distribution.

In 2006, UBC Press suspended its agreement with Baker & Taylor¹⁹ in order to partner exclusively for one year with Canadian Electronic Library, an ebrary partner offered by Gibson Library Connections and affiliated with Gibson Publishing Connections, the Canadian operator of UBC Press's DAMD service

¹⁸ Peter Milroy (director of UBC Press), personal communication.

¹⁹ That is, UBC Press did not grant Baker & Taylor the rights to distribute 2006 titles; B&T still distributes past UBC Press titles.

(see section 1.5.2.4). Through this agreement, UBC Press ebooks are currently offered only in PDF format,²⁰ in a perpetual access, single viewing model, to CEL's library affiliates. In order to establish a presence as the foremost distributor of Canadian scholarly ebooks, CEL offered to scan and generate PDFs from hard copy archives for UBC Press's entire backlist without immediate charge (the \$90 PDF creation fee to be instead deducted from royalties) in exchange for this year of exclusivity. This short respite from multiple avenues of distribution has provided a perfect period for integrating new strategies for addressing distributor requirements, in anticipation of fall 2007. After August 2007, UBC Press will resume sales through Baker & Taylor, in addition to its current distributors – CEL, ebrary, and the non-competitive service of Google Book Search – as well as pursue new ebook distributors, such as MyiLibrary, with which UBC Press is in negotiations. As will become clear in the following sections, the significant amount of work involved in meeting the needs of only three avenues of ebook distribution has been considerably streamlined, suggesting the valuable place of digital asset management/distribution service in a typical multi-distributor scholarly ebook production process.

²⁰ An exception is its Globalization and Autonomy Series e-compedium, which is an XML-based, external companion website to UBC Press's Globalization and Autonomy book series. This type of digital book, however, has few implications for UBC Press's ebook production process, as it has been produced out-of-house by the editors of the series.

2.2 UBC Press and Digital Asset Management/Distribution

UBC Press engaged codeMantra's DAMD²¹ services in 2006 in an effort to streamline its ebook production and digital asset management. CodeMantra's DAMD service is not without fault; its implicit reinforcement of a multiformat ebook market aside, codeMantra's DAMD has many operational limitations. However, codeMantra's services have nonetheless provided UBC Press ebook production with a major functional boost. According to codeMantra's website, the full services available to publishers through the company are:

- Ebook creation
- Composition and typesetting
- Technical editing and copyediting
- Data conversion and processing
- Abstracting, indexing, and coding
- Image processing
- Content management and delivery
- Database development
- Custom content solutions

Of these, the most important for the ebook production process (and the only services utilized by UBC Press) are ebook creation, and content management and delivery.

²¹ All material regarding codeMantra's services has been gathered from their website, <http://www.codemantra.com>, UBC Press's Collection Point site, <http://ubc.gibsonpublishingconnections.com>, and personal experience.

2.2.1 Ebook Creation Services

CodeMantra uses XML-based software to convert files uploaded from UBC Press staff or designates (freelancers) into its Universal PDF format, which is offered at a blanket rate of \$90 USD per title. CodeMantra conversion services are also available for Handheld Suite (MobiPocket, MS Reader/.LIT and Palm) and Print on Demand format; at present, UBC Press uses neither of these services.

2.2.1.1 Universal PDF

CodeMantra's primary ebook creation service for UBC Press is the translation of a title's component files into what codeMantra calls a Universal PDF (uPDF) file. As a response to the logistical problem of coordinating e-distributor submission requirements, the uPDF unifies many of the major distributors' needs, in what codeMantra describes as a "fully compliant file created for nine of the leading electronic book distributors, as well as several other sales channels such as Amazon's Search Inside and Google Book Search" (for a complete list of distributors accepting the uPDF, see Appendix B).

Rather than produce individual PDF versions of its ebooks for each distributor, a process that requires additional manpower and specific technological skill sets, UBC Press outsources to codeMantra the creation of a

single file that will be accepted by all of its distributors. The uPDF is an Adobe PDF document, as opposed to a unique format; its benefit lies not in its technology but rather in its specifications. According to codeMantra's website, the specifications of the uPDF that ensure general e-distributor acceptance are:

- Properly embedded fonts
- Cropped registration marks and blank areas
- Image and text disclaimers
- Bookmarked table of contents
- Cleaned-up and inserted cover pages
- Linked table of contents (to three levels)
- Linked endnotes
- Headings linked to table of contents (to three levels)
- External URLs
- Images optimized for Web delivery
- Indices linked (optional service)
- Complimentary file delivery

Nothing prevents UBC Press from producing files that match these specifications itself, except for the time and technological understanding involved in integrating them. However, the cost of the requisite time and expert labour is at present both prohibitive and higher than codeMantra's uPDF creation fee (see section 2.3 for a more detailed financial summary).

CodeMantra creates uPDF files for UBC Press from the publisher's source files, whether those are low-resolution PDFs or design files such as Adobe InDesign or PageMaker. CodeMantra has provided UBC Press with similar

documents for ebooks scanned from hard copy (per the Press's backlist scanning agreement with CEL); these documents are termed Universal Image PDF (uiPDF) files,²² and they are accepted by only six distributors (see Appendix B). If distributing these uiPDF titles through all possible organizations proves important to the Press, codeMantra can convert them into uPDFs at an additional cost of \$2 per page; however, because UBC Press's current distributors accept the uiPDF format, it has declined this service. Both PDF formats represent value-added ebooks: not only do they conform to distributor specifications, but codeMantra also performs the labour involved in introducing extra features such as linked indexes, features that are not required by e-distributors and would otherwise need to be incorporated by trained staff or freelancers.

2.2.1.2 Print on Demand

Print on Demand (POD), while not traditionally a component of ebook production, is quickly becoming an important consideration for UBC Press's ebook production process. As with e-distributors' needs, UBC Press's production department must factor POD requirements into the creation of digital formats; like e-distributors, POD service providers have specific file requirements. If UBC

²² The key technological difference between uPDF and uiPDF is that uiPDF is raster-based (meaning that its pages are treated as pictures composed of pixels), unlike the vector-based uPDF (in which pages are "drawn" from the relationships between points, which allows Adobe Acrobat to recognize and manipulate text).

Press pursues POD (with scholarly publishing's typically small print runs and often unpredictable reprint needs, this is not unlikely), it can bundle the creation of codeMantra's POD-ready PDF format (accepted by most POD services) with uPDF creation, at a per-book fixed cost of \$25 (\$50/book when done separately).

2.2.2 Collection Point

As outlined in section 1.6, an ebook is much more than a single file; rather, for any given title, UBC Press must gather multiple component files and produce multiple end-use formats. A typical UBC Press ebook consists of a uPDF file, which is generated from two or more low-resolution PDF cover and text files, files that are themselves created from PageMaker or InDesign files (for both the cover and the text) that unite multiple image and text files. When these digital needs are combined with print requirements (high-resolution PDFs, plate insert files, etc.) and multiplied across the Press's entire list, effective management of a title's digital files becomes vital for the creation of any book, electronic or otherwise. Management of a book's various digital files must occur at every stage of its life; throughout a title's editorial shaping, distribution, and archiving, file management becomes an overarching feature of production. The publishing industry as a whole must commit to long-term solutions to format divergence (solutions like developing and implementing digital standards), but before such

solutions are in place (and afterward as well), effective file management is an indispensable factor of ebook production. Yet despite the increasing need to systematize and regulate digital assets, many publishers treat their files in a haphazard manner. Publishers frequently volley a book's files, in multiple versions and iterations, among staff, freelancers, and distributors with little if any tracking. Digital assets may be scattered across in-house network servers, removable drives, disks, CDs, and individual desktops and email inboxes, where they are prone to misidentification or retrieval problems at best, and loss at worst.

As long as the reality of multiple file requirements exists for publishers, DAMDs are an effective means of organizing all of the files involved in the life of a title because they facilitate file identification, transfer, tracking, and storage. CodeMantra's DAMD interface, Collection Point, is an external server with a Web interface, to which publishers can upload files, files that can then be organized, stored for future use (e.g. reprints or new editions), and/or transferred between parties involved in production and distribution. Collection Point allows UBC Press to identify its books and all associated files; transmit files between individuals such as Press staff, freelancers, and distributors; analyze production trends through auto-generated data reports; and maintain an archive. Although the Collection Point interface is by no means a digital ebook solution, and indeed

reinforces the variegation of the present ebook market, the streamlining of the ebook publishing process that it has provided at UBC Press has been marked.²³

2.2.2.1 File Identification

The vast array of works in UBC Press's list, as well as the numerous files assigned to each, require specific methods of identification if they are to be usefully manipulated and employed. Collection Point provides this by assigning a unique record to each individual work, within which all of its various composite files are placed. These records are made distinct and identifiable through the assignment of metadata and content files.

Metadata:

The blueprint of a title is its metadata, or the data that describes its data (more simply, the facts about itself). Within a publishing house, titles are identified by elements such as title, author, ISBN, or copyright year. Within Collection Point, each unique title record is created through the assignment of that title's specific metadata, allowing UBC Press to organize and manipulate works within the

²³ Based on personal communications with Anna Friedlander (UBC Press Editor, and UBC Press's Collection Point Administrator), Darcy Cullen (UBC Press Editor), Holly Keller (UBC Press Assistant Director, Production and Editorial Services), and Peter Milroy (UBC Press Director).

system based on several fields. Collection Point's list of available metadata fields is extensive:

- Publisher Name
- File Name
- ISBN
- Publication Date
- Print List Price
- Print Currency Type
- ISBN 13
- ePublication Date
- Ebook List Price
- Ebook Currency Type
- Title
- Author
- Contributor 1
- Contributor 1 Role
- Contributor 2
- Contributor 2 Role
- Contributor 3
- Contributor 3 Role
- Publisher Name
- Publication Location
- Language
- Distribution Type
- View Type
- Batch #
- Collection Code
- BISAC1
- BISAC2
- BISAC3
- Subject
- Title URL
- Format
- File Size (in MB)

At UBC Press, the metadata assigned to each book currently are²⁴:

- Publisher Name
- eISBN 13
- Author Name
- Title
- Copyright Year
- Ebook List Price
- Format
- ISBN 10
- Author Role
- File Size
- Language
- Print List Price

²⁴ This list of metadata is preliminary. UBC Press is in the process of integrating a Klopotek database with its operations, at which point the Press will incorporate full ONIX metadata sets into each title record. See the following note for an explanation of ONIX metadata.

Metadata serve as the identifying features of a title record, allowing publishers to know what they are working with when they use Collection Point. When staff members wish to work with certain titles, they can search for them using any of their differentiating metadata (i.e., ISBN, title, author, copyright year). They can also use this metadata for summaries or interpretations of the data contained within the system (for example, a summary of ebook sales for all 2007 titles) by generating reports (see section 2.2.2.3 for more information on reports). Metadata are equally important for organizing and identifying individual titles outside of the Collection Point system, particularly during printing and distribution. Titles' accompanying metadata are used by distributors, aggregators, wholesalers, printers, libraries, and end users as important tracking and organizing devices.

Metadata are uploaded from the publisher to Collection Point using a simple Excel spreadsheet or an ONIX-BISAC²⁵ file containing all of the fields that the publisher wishes to include. UBC Press currently adds metadata to Collection

²⁵ ONIX, or Online Information Exchange, is a standard XML format used by publishers to communicate metadata; see Book Study Information Group, "ONIX (Online Information Exchange)," *Book Study Information Group* http://www.bisg.org/onix/onix_faq.html (accessed August 11, 2007). BISAC, or Book Industry Standards and Communications, refers to standard metadata types. See Book Study Information Group, "BISAC Metadata Committee," *Book Study Information Group* <http://www.bisg.org/bisac/metadata/index.html> (accessed August 11, 2007) Requirements gathered from Gibson Publishing Connections, "Introduction to Collection Point for Canadian Publishers," unpublished informational booklet for publishers, 5.

Point in an Excel file, but intends to use ONIX once a new press-wide Klopotek database is in place that will facilitate the production of ONIX metadata.

Collection Point notifies the Press of any problems with its metadata submission through an auto-generated message on the metadata upload screen. Once metadata have been uploaded for a book, Collection Point automatically creates a record, which is then available for content file uploads.

Content Files:

Once a record is created, Press staff members and freelancers upload the book's constitutive files. Within individual title records, UBC Press organizes files according to set formats. The Press is able to define the file formats it wishes to make available for upload; at present, UBC Press has elected to include the following:

- CIP page for paperback InDesign
- CIP page for paperback PageMaker
- CIP page for paperback PDF
- CIP page for paperback QuarkXPress
- Cover files zipped
- Cover files for later paperback zipped
- Cover front JPG
- Cover front PDF
- Cover front TIFF
- Cover info form for printer
- Ebook Universal PDF
- Ebook Universal Image PDF
- Full book PDF from codeMantra scan
- Manuscript files
- Plates insert for printer PDF hi res
- Plates insert for printer InDesign
- Plates insert for printer PageMaker
- Plates insert for printer QuarkXPress
- Project files archive

- Replacement pages first reprint
- Replacement pages second reprint
- Temp file transfers
- Text Complete InDesign
- Text Complete PageMaker
- Text Complete PDF final low res
- Text Complete PDF for printer hi res
- Text Complete QuarkXPress
- Text info form for printers
- Uncorrected text proof.

By defining specific upload formats, the Press can ensure that each title has the files, and in the correct formats, necessary for uPDF creation, printing, and distribution. By designating file types, UBC Press has also imposed a clear file nomenclature. Staff members do not need to search within several PDF files to determine which is the high-resolution version required by the Press's printer; rather, they will simply look for the file uploaded to that field.

Within each title's record, Press staff members can easily survey its content files. All format types possible for each title (following the above list) are displayed within each record; format fields with files attached appear in blue, while format fields with no attached files appear in black. UBC Press staff members can quickly survey a title's upload page in order to determine if specific content files (such as cover files) have been added to the Collection Point system.

Because each file format type can be occupied by only one file at a time, staff members are able to access only the most recent version of that format type.

As a result, they are less likely to engage in redundant work caused by the confusion of file versions. However, because files within a specific format type are replaced with newer uploads, and because nothing in the Collection Point system will prevent staff members from uploading incorrect files, users must be extremely careful when uploading. Files overwritten by new uploads are not irretrievable; if publishers contact codeMantra, older versions can be resurrected. Still, the inconvenience of resurrecting mistakenly overwritten files necessitates caution.

2.2.2.2 File Transfer

Collection Point acts as a hub for all parties involved in the creation and distribution of a UBC Press title. Rather than risk the loss or misidentification of files volleyed between staff, freelancers, and distributors, Collection Point has allowed the Press to streamline and monitor a title's movement. Collection Point facilitates transfer of digital assets through file uploads and file delivery.

Uploads and Local Downloads:

The production of a title necessitates the combined work of many individuals. At UBC Press, in-house work on a manuscript can be facilitated by file storage on the network server; however, although UBC Press staff are involved in many of the production functions required to create a complete title,

most such functions are outsourced to expert freelancers.²⁶ Collection Point's upload and local download features allow staff to coordinate the transfer of files between these freelancers and staff. Rather than transmit files through email, which risks loss and confusion, staff members and freelancers place files directly within a title's record. In the development of a typical title, staff members upload the text files of a manuscript, which a copy editor then downloads, works with, and re-uploads. After editing is complete, a typesetter then downloads final text files for the typesetting process, and uploads page proofs as PageMaker/InDesign, low-resolution PDF, and high-resolution PDF formats. Similarly, a cover designer submits completed cover designs directly to Collection Point, 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 rather than other means of file transfer, such as email, all parties involved in the production of a title ensure that they are using the correct and most recent file. Files also become available for transfer to staff members in other departments of the Press; for example, page proofs and cover files stored in Collection Point are used by marketing staff.

²⁶ See Megan Brand, *Outsourcing Academia: How Freelancers Facilitate the Scholarly Publishing Process* (MA Thesis, Simon Fraser University, 2005).

UBC Press has developed guidelines for freelancer uploads (which components are required, and how they should be packaged – see Appendix C). Freelancer uploads to Collection Point have already made the transfer of working cover and text files far less likely to result in confusion or loss. Centralizing the transmission of manuscript files, page proofs, cover design, and final book files has streamlined the production process, eliminating redundancy, file duplication, and the storage of multiple versions. Staff no longer need to monitor freelancer submissions within the UBC Press server, nor must they navigate through incorrect format types or unclear file labeling.

Of course, when granting out-of-house individuals access to key files, security measures to protect Press materials are vital. UBC Press's Collection Point administrator, Anna Friedlander, maintains security by allowing password-protected access to the system to designated individuals. Similarly, she is able to define which functions designated users are allowed to perform; currently, available functions include searching, generating reports, and distributing and uploading files.

Files destined for users who are not designated users of UBC Press's Collection Point site are sent to them via a "shopping cart" interface similar to those used on online shopping sites. Using this shopping cart, staff members

select specific files within a title's record for transfer, files which they can then send to a distribution partner, a third party, or codeMantra for uPDF conversion.

Sending to a Distribution Partner:

Collection Point allows publishers to designate more than a dozen distribution partners (in Collection Point, this means organizations with which they have longstanding distribution agreements). At UBC Press, those distribution partners are Friesens (the press's printer); Gibson Library Connections; Gibson Library Connections at ebrary (CEL); Google Book Search; and Library and Archives Canada e-book legal deposit. Each distributor designated as a partner is associated with their specific file requirements, and deliveries made to them automatically include ebook uPDFs and correct metadata. All major ebook distributors, including Google, Amazon, Microsoft, NetLibrary, ebrary, and more, can be set up as distribution partners.

Sending to a Third Party via FTP:

Collection Point's "send to third party via FTP" feature allows UBC Press to send files via FTP²⁷ to individuals or services other than distribution partners or authorized Collection Point users, such as reviewers. Third parties need not have access to the UBC Press Collection Point site; this file delivery takes place

²⁷ That is, *File Transfer Protocol*; in which files are uploaded to a site and made available for the site's users to download.

through a separate, individualized FTP site. Through this delivery mode, Press staff members indicate a destination email address, and the Collection Point system sends an email notifying the recipient of a password for a temporary dedicated FTP URL from which they can download the specified material within 72 hours. Files sent by this method of delivery can be made only temporarily usable, which allows UBC Press more control over the use of its files by third parties.²⁸

Sending to CodeMantra:

Once a title's component files are finalized, uploaded, and ready for e-distribution, UBC Press is then able to send the title to codeMantra for conversion into a uPDF. In order to request conversions, Press staff members place the title's low-resolution PDF file and cover file into their shopping cart and select the "send to codeMantra for file conversion" option. CodeMantra generates the uPDF file and uploads it to the title's record, at which point it is available for delivery to distributors.

2.2.2.3 File Tracking

In addition to file identification, UBC Press uses the metadata and content files uploaded to Collection Point to summarize and evaluate file activity, as well

²⁸ Bob Gibson, personal communication, July 20, 2007.

as to monitor the Press's interaction with the system and outside partners.

Through Collection Point's Report feature, staff members generate summaries of the UBC Press Collection Point system's status, descriptions of which follow.

All Titles and Full Metadata:

UBC Press is able to generate charts listing all of the individual title files that have been uploaded to the system, as well as full metadata for each. These reports allow staff to survey which titles and files have been added to the system, and which have yet to be uploaded. These summaries have proved most useful during the Collection Point inception process, when the Press moved its production activities to the system.

If staff members do not need a full metadata report but rather summaries based on select metadata, they can generate specific metadata reports. A typical example would be a report based solely on titles published in 2006.

Activity Tracking:

Activity tracking generates reports based on user activity; individual users can monitor their actions in UBC Press's Collection Point site in order to keep track of what work they have already done and what remains. The Collection Point administrator (Anna Friedlander) can track the activities of all authorized users; for example, she might determine whether a freelancer has uploaded a

specific file. The remaining report types (title history and summary) are similarly available to the system administrator (although codeMantra might be well advised to make them available to other system users).

Title History:

A title history report is similar to an activity report, except that rather than organize actions by user, it displays what actions have occurred for specific titles. A typical title history report will include which file formats have been uploaded and transferred, by whom, and when. This type of report is useful when UBC Press wants to monitor activity relating to individual titles.

Summary

Summary reports are among the most useful for UBC Press. Through this type of report, staff can evaluate the activity of specific formats and distributors (summaries of all titles and all users are also available, but these closely parallel other report types already discussed). A format summary displays which titles have files in specific formats, allowing UBC Press to see at a glance which titles need, for example, a uPDF file. A distributor summary indicates which titles have been sent to each distribution partner, which can help UBC Press determine whether or not specific distributors are an asset.

2.2.2.4 Archiving

Neither good production nor good management ends when a book is sent to distributors. Publishers must maintain robust archives, not only for posterity but also, more practically, because of the possibility of reprints or subsequent editions. UBC Press has maintained archives since the late 1980s, transitioning from removable hard drives kept in locked storage to CDs, and finally, to Collection Point. UBC Press is currently in the process of moving CD-ROM archive files to the codeMantra server.

Archiving through Collection Point offers UBC Press advantages over its past in-house archiving practices in terms of both convenience and security. Archiving through Collection Point occurs concurrently with UBC Press's production of a title, as the most recent content files uploaded to a title's record remain there indefinitely. File retrieval occurs as a local download, obviating the need to search through physical files or connect with storage devices.

The possibility of damage to archives stored in-house on disks, drives, or CDs, as well as potential file inaccessibility due to technology changes, give Collection Point archiving significantly more security than UBC Press's past

archiving methods. CodeMantra ensures the safety of files through server redundancy: it maintains files on triplicate servers in the US and India.²⁹

2.2.2.5 Collection Point Summary and Suggestions

No program is perfect, just as no process is perfect. Regardless of the long-term implications of DAMDs, on the level of function, codeMantra could improve the ergonomics of its interface in multiple ways. For example, following activities like uploads, users are rarely returned to their last page, but rather to start pages. CodeMantra could also improve system security by allowing administrators to authorize users to perform only specific format type uploads, thus avoiding potential misplaced or overwritten files. However, these areas for improvement are minor when contrasted with the significant clarifying and organizing function that this system, as well as DAMDs in general, provides to ebook producers.

2.3 Financial Summary of Ebook Production at UBC Press

The immediate organizational benefits of DAMDs become even more appealing for ebook producers when one considers the expenses that publishers

²⁹ Although technological change resulting in access problems remains a possibility, codeMantra is far less likely to barricade its own access to server files than is UBC Press (in an example common to publishers, through the purchase of new IT systems that do not support past removable drive devices containing archive files).

would need to incur in order to perform similar functions in-house. UBC Press's transactions with codeMantra, outlined below, demonstrate the extent to which codeMantra's economy of scale (the ability to reduce the per-unit cost of infrastructure, such as technology and skilled employees by producing products in bulk) makes in-house DAMD development unreasonable for publishers of UBC Press's stature

2.3.1 Ebook Creation

The codeMantra ebook creation services used by UBC Press (i.e., uPDF only) are offered at a flat rate of \$90 per book. Therefore, the total yearly cost of uPDF creation multiplied across UBC Press's entire list (on average, 60 titles per year) is approximately \$5,400.

When compared to the costs of submitting regular low-resolution PDFs to distributors, with an approximate cost of \$100 USD per distributor-performed conversion,³⁰ PDF submission at UBC Press (after its exclusivity arrangement with CEL expires, allowing distribution through CEL/ebrary, Baker & Taylor, and MyiLibrary) would cost \$300 USD per title, for a total estimated yearly cost of \$18,000 USD (approximately \$18,900 CAD).

³⁰ Averaged according to figures listed by NetLibrary, ebrary, MyiLibrary, and Baker & Taylor on their websites (<http://www.netlibrary.com>, <http://www.ebrary.com>, <http://www.myilibrary.com>, and <http://btol.com>, respectively; all accessed July 19, 2007).

Alternatively, the cost of the labour hours required to produce tailored PDFs for each distributor in-house, including training and performing the tasks of linking tables of contents, endnotes, and occasionally indexes, would approach \$200 CAD per title (approximated from average freelance typesetter and designer hourly rates for an estimated five hours of work), or \$12,000 CAD across UBC Press's yearly list.

2.3.2 Collection Point

Collection Point pricing operates on a per-title basis, not per file or per megabyte. At UBC Press, Collection Point costs equal roughly \$5,000/year. The labour hours saved by Collection Point, and the labour and IT infrastructure costs required to provide commensurate functions (including the costs of a dedicated website and triplicate archives) are difficult to calculate, but they are safely in excess of the \$5,000 Collection Point fee.

2.3.3 Revenue

UBC Press's ebook revenues are currently modest, but sufficient to warrant the effort involved.³¹ As ebooks gain popularity and libraries pursue digitization projects, and if ebook formats and file requirements become more

³¹ Due to confidentiality agreements with distributors, royalty structures and revenues cannot be reproduced here.

divergent, the importance of minimizing ebook production costs will become only more apparent.

2.4 Ebook Production at UBC Press: Conclusions

The codeMantra uPDF creation and DAMD services have proven enormously helpful in UBC Press's efforts to produce and distribute ebooks. They have made the management of files and distributor needs an orderly, automatic, and accessible process. In addition to their usefulness, the services provided by codeMantra save both cost and labour, allowing UBC Press staff to concentrate their efforts on producing quality scholarly works. In the face of a muddled ebook market, the adoption of a DAMD has proven an effective and necessary means of streamlining ebook production.

Yet, however well external digital asset services might address a significant problem in the ebook industry, they do very little to remove that problem in the long term. In fact, they are more likely to institutionalize those problems and assign the task of addressing them to publishers. Rather than induce distributors to simplify their rules for ebook submission, the creation of a uPDF by an external service provider makes meeting distributor requirements a publisher responsibility, one that is mediated and arcane, and makes staff less

aware of ebook requirements. Similarly, the use of a DAMD obscures the fact that current technologies and file format specifications are convoluted.

CodeMantra's services have significantly simplified UBC Press's ebook production process. But what logically follows is a coordinated industry effort to simplify ebook production in general, ensuring that what should be a stopgap solution does not become a fixture of ebook publishing.

3. STANDARDIZATION INITIATIVES

Before the need to conform to multiple format requirements becomes entrenched, publishers must commit to finding long-term solutions. However, for most scholarly publishers (UBC Press included), finding long-term solutions is a lower priority than addressing the constant challenges of running a publishing house. Indeed, the challenges of maintaining a scholarly publishing operation are so great that few academic presses have the financial resources or the industry clout to bring about widespread change. For this reason, ebook producers look to external digitization initiatives for change. Developing a sustainable solution to ebook format needs depends on the establishment and widespread acceptance of open standards, creative DRM, and publisher-library ebook cooperatives. While external, these initiatives cannot succeed in isolation; they depend on publishers to cooperate with their goals.

3.1 International Digital Publishing Forum: The Open Publication Structure Specification

As indicated in section 1.4 on Digital Rights Management, the longevity of the ebook market requires widespread adoption of open standards for ebook

formats, which allow unrestricted use and interoperability across multiple platforms.³² One such open standard, which is growing in importance for ML ebook production, is the *Open Publication Structure Specification* (OPS). A project of the International Digital Publication Forum (IDPF), a trade and standards organization concerned with developing and promoting e-publishing, OPS is a specific set of XML encoding standards. OPS is extensive; however, publishers intending to produce OPS ebooks need not be expert in markup to employ it, as Adobe Indesign CS3 supports the direct generation of OPS documents through an export function.

By making their ebooks and reading devices OPS compliant, ebook and technology providers ensure that their products will be interoperable with the wide variety of ebooks, systems, and devices that end-users might possess; that technology change will not render ebooks inaccessible and devices obsolete; and that the ebook enterprise will involve less conversion and coordination effort.³³

³² It is important to make the distinction between open standards and open access. *Open standards* refer to technologies that are interoperable and manipulable, while *open access* refers to content that is offered to consumers free of charge.

³³ For more information on IDPF and OPS, see the IDPF website at <http://www.idpf.org>.

3.2 Open Standard PDF

For publishers whose works are not appropriate for flexible and reflowable design, the PDF is already a viable standard format.³⁴ The PDF is widely accepted as an industry basic, and the universality of its use indicates the benefits of standardization. As an additional step towards complete standardization, on January 29 of this year Adobe made its full PDF 1.7 specification (the version created with Adobe's most recent Acrobat 8) open format. Many components of PDF were open format prior to this development, but in cooperation with the Association for Information and Image Management, the PDF format in its entirety has now been submitted for ISO registration.³⁵ If it is approved (as it likely will be), PDF will be recognized as an international standard as ISO 32000. PDF has been a common feature of publishing (ebook and otherwise) for many years; however, acceptance as an international standard means that conflicting iterations of PDF will be reduced. As Melissa Webster, IDC's Content & Digital Media Technologies program director asserts, "when...distributors take a published spec [the PDF] and then add to it, they

³⁴ See Adobe, "Products and Services," *Adobe*, http://www.adobe.com/products/#acrobat_family (accessed July 30, 2007) for more information.

³⁵ See Adobe, "PDF Reference," *PDF Developer Center*, http://www.adobe.com/devnet/pdf/pdf_reference.html (accessed August 8, 2007); see also Bill McCoy, "PDF Slated to Become ISO Standard, 'All Inclusive,'" *Bill McCoy*, <http://blogs.adobe.com/billmccoy/> (accessed August 1, 2007).

may feel they are 'improving' it...but if they incorporate their changes to it in a series of products that become a 'de facto' standard, it creates a conflicting version of that spec."³⁶ As an ISO standard, parties such as ebook distributors can have input in the specifications of PDF, reducing their need to develop proprietary versions.

Adobe has also begun working with IDPF to integrate OPS-compliant XHTML ebooks with its new Adobe Digital Editions product, an ebook digital bookshelf and reading device. This step ensures that Digital Editions supports both of the major ebook open formats, the ML-based OPS and the ISO 32000 PDF. By cooperating with IDPF, Adobe has made a significant step towards ensuring that whether publishers or end-users need flexible ML or design-controlled PDF ebooks, they will be able to use them in an open standard platform.³⁷

3.3 Digital Rights Management Initiatives

Standardization of open formats will not significantly improve the ebook market unless it is paired with new approaches to DRM. Even if ebook

³⁶ Quoted in Don Fluckinger, "Adobe, AIIM Make ISO Push for PDF," *PDFZone*, <http://www.PDFzone.com/article2/0,1895,2088267,00.asp> (accessed August 1, 2007).

³⁷ The above was informed largely by Bill McCoy, "Adobe Delivers Digital Editions 1.0," *Bill McCoy*, http://blogs.adobe.com/billmccoy/2007/06/adobe_delivers.html (accessed August 2, 2007) and McCoy, "PDF Slated to Become ISO Standard."

stakeholders fully standardize their formats, as long as they modify those standard formats with custom DRM, ebooks will maintain their current divergence. Both IDPF and Adobe are developing or have developed DRM standards to accompany their formats; the IDPF has established a "Rights and Rules Working Group" intended to create an IDPF DRM specification, and Adobe DRM capabilities are a blanket provision with Adobe Acrobat and Reader. However, even if IDPF develops a standard and ebook providers widely adopt both it and Adobe DRM, DRM will continue to present a problem for ebook producers. DRM is by nature a closed format, as making DRM a universal standard makes it vulnerable to sabotage; therefore, in order for piracy to be prevented, ebooks files are customized with closed, proprietary rights technologies. In order to simultaneously protect ebooks from unauthorized use and facilitate universal and durable accessibility, ebook stakeholders need to develop new DRM paradigms. Instead of locking ebooks themselves within proprietary formats, ebooks can be protected from illegal use through what might be thought of as gatekeeping practices. Gatekeeping focuses on securing the transfer of ebooks between provider and consumer. Tethered access is one such gatekeeping practice, although at present most tethered access distributors (most notably ebrary) utilize proprietary formats (such as EDF); again, this type of DRM must be combined with open formats (for example, through OPS-

compliant tethered access) if conflicting format types are to be reduced. Another gatekeeping practice (one that is generally paired with tethered access) is user authentication, which involves both password protection and technical mechanisms that identify users' computers. Over the long term, ebook providers will need to pair gatekeeping practices with other DRM paradigms such as "social DRM." The Association of American Publishers suggests that DRM is not merely a technological issue, but rather one that involves social and legal actions. The AAP suggests that ebook providers combine education concerning legal use with superior service for legal users, such as customized bookshelves on reading devices that organize users' entire ebook collections (much like iTunes organizes music purchases), and tailored ebook purchase packages offering discounts and recommendations.³⁸ On his Adobe blog, Bill McCoy also highlights the use of social DRM by The Pragmatic Programmers' in the form of digital "For the Exclusive Use of ..." stamping.³⁹ This type of creative DRM, as well as gatekeeping practices, represent the type of rights protection that will prove most successful for discouraging illegal use without hindering accessibility and operability entirely.

³⁸ Association of American Publishers, *Digital Rights Management for Ebooks*, 17.

³⁹ See Bill McCoy, "Steve Jobs: 'Eliminate Music DRM!': So, What About ebooks?" *Bill McCoy*, http://blogs.adobe.com/billmccoy/2007/02/steve_jobs_elim.html (accessed August 1, 2007).

3.4 Association of Canadian Publishers

The development of approaches and standards designed to promote a uniform and interoperable ebook market will have little effect if pressure to adopt them is not placed on distributors and technology providers. Scholarly publishers can best exert this pressure by partnering with libraries and other publishers, in an effort to reduce dependence on industry middlemen.

An initiative that identifies means of doing so has recently been undertaken by the Association of Canadian Publishers (ACP), an industry-member organization that aims to promote the activities of publishers in Canada. In her report for the ACP on digitization in publishing, Diane Davy correctly identifies the problem facing ebooks (too many files and formats) and the interim solution (DAMs), and suggests that publisher and library consortia move towards standardizing ebook formats and distribution.

Davy's report is focused on "work[ing] toward streamlining and digitizing...internal workflow processes."⁴⁰ Davy argues that, while most of the industry's attention is fixed on enhancing revenue through the sale of digital content, it is in fact through streamlining digitization processes that publishers

⁴⁰ Diane Davy, "The Impact of Digitization on the Publishing Industry," report prepared for the Association of Canadian Publishers, March 31 (Ottawa: ACP, 2007), 4.

will receive the greatest financial gains (a sentiment echoed by Adobe representative Bill McCoy).⁴¹ For the most part, Davy focuses on DAMDs as a means of streamlining. However, she also identifies an important step toward the long-term streamlining of ebook production in her suggestion that publishers or public organizations serving the publishing industry form a cooperative designed to distribute ebooks to libraries directly. Presumably, such a cooperative would present a large enough catalogue of titles to libraries to be a viable competitor, and it would simultaneously eliminate distributor royalty deductions and the need for tailored PDF submissions. This type of cooperative inherently requires open formats, if publishers hope to elicit widespread compatibility with libraries' divergent platforms. This suggestion is, however, merely that; the logistics of forming such a cooperative are not discussed.⁴² It remains to be seen whether the ACP or publishers will act on this suggestion.

3.5 Canadian Research Knowledge Network's Digital Content Infrastructure

In the absence of a publisher consortium, Canadian university libraries have developed a cooperative, the Canadian Research Knowledge Network (CRKN), which is "a partnership of Canadian universities, dedicated to

⁴¹ Ibid., 10.

⁴² Ibid., 18.

expanding digital content for the academic research enterprise in Canada.”⁴³

Since 2004, this not-for-profit organization composed of 72 university member libraries has developed Canada’s digital research infrastructure, emphasizing the licensing of electronic versions of scholarly journals and databases.

CRKN has recently announced its “Digital Content Infrastructure for the Social Sciences and Humanities” project (DCI), which will progress through 2009. This approximately \$48 million initiative (\$20 million in funding comes from the Canadian Foundation for Innovation, with the remaining funds coming directly from CRKN partner libraries) is intended to “secure durable rights to a cohesive portfolio of digital content, in multiple languages and multiple formats, in the Human and Social Sciences (HSS) for researchers at 67 participating Canadian universities.”⁴⁴

CRKN is currently in the process of soliciting proposals from ebook distributors for the purchase of content. Proposals are to be weighted according to publishers’ ebook capability (including their amenability to open standards and interoperable DRM), technical support services, quality, licensing approach,

⁴³ Canadian Research Knowledge Network, “What is the CRKN?” *CRKN/RCDR Canadian Research Knowledge Network*, <http://researchknowledge.ca/en/index.jsp> (accessed August 1, 2007).

⁴⁴ CRKN, “Digital Content Infrastructure for the Human and Social Sciences (DCI) Project Implementation Process: Frequently Asked Questions (FAQ)” *CRKN/RCDR Canadian Research Knowledge Network*, <http://researchknowledge.ca/en/programs/documents/DCIConsultationFAQ-English.PDF> (accessed August 2, 2007), 1.

usage reporting, access models, and cost.⁴⁵ It is therefore incumbent upon publishers to consider their capabilities in these areas when producing ebooks. The primary purpose of CRKN is not to establish industry ebook standards, and it offers no aid to publishers in this regard, but CRKN's DCI does offer publishers a financial impetus for rapidly develop of best ebook production practices, and on an open format model. CRKN implicitly recognizes the need for interoperable standards, rather than proprietary formats, to unite significant industry stakeholders and to address their needs as a cohesive whole.

3.6 Standardization Conclusions

Although CRKN's DCI is still under way and its success or lack thereof has yet to be determined, it provides an example that publishers and organizations like the ACP could follow. By supporting cooperatives that represent their mutual interests, publishers could either remove the need for distributors and their associated requirements or, as a cohesive whole, exert enough pressure on distributors to encourage standardization under open formats and interoperable DRM. Even though cooperation on these types of initiatives represents significant effort on the part of publishers, the long-term

⁴⁵ CRKN, "Electronic Content Expansion Project: Request for Proposal," informational package sent to publishers (Ottawa: CRKN, September 12, 2005), 8, 16.

effort and costs involved in coordinating multiple formats and files are considerably higher.

4. CONCLUSIONS

The current ebook market is characterized by convolution. Part One of this report outlined the role that multiple file formats, reading devices, methods of access, Digital Rights Management, and distributor requirements play in this complicated market, as well as the impact they have on the ebook production process.

Part Two progressed to describe the actions that UBC Press has taken to streamline its ebook production in the face of the industry's multiplicity. Specifically, it described UBC Press's adoption of ebook creation and digital asset management/distribution services, provided by codeMantra.

Part Three discussed standardization initiatives capable of providing a solution to the manifold needs of the ebook market. Open standards (such as International Digital Publication Forum's Open Publication Structure Specification or open standard PDF), non-proprietary digital rights management, and library/publisher distribution cooperatives present viable means of reducing the format variance that currently stifles the ebook industry.

In Andersen Consulting's 2000 ebook study for the Association of American Publishers,⁴⁶ its overly optimistic predictions for the growth of the ebook market were tempered by warnings of the market disorganization and restriction that could result from publishers' failure to establish universal standards for the dissemination of ebooks. This prediction, at least, appears to have come true. As publishers like UBC Press seek to address the confusion in the ebook market and its impact on ebook production, the solution has thus far been intermediation. In order to allow a simple, streamlined production process while still accommodating distributor and end-user needs, third-party DAMD providers have become necessary for day-to-day operations. At UBC Press, codeMantra's provision of these DAMD services has proved a very effective means of addressing the divergent formats, specifications, and requirements involved in ebook production.

Yet the necessity and effectiveness of DAMDs in the face of the messiness of the ebook market do not remove that messiness. Confusion of formats and digitization standards will continue to stunt the potential of the ebook market as long as that confusion is present. It therefore seems wise for publishers of ebooks to support standardization of the industry as a whole. The development of a streamlined, interoperable ebook market is vital for the long-term success of

⁴⁶ Andersen Consulting, *Reading in the New Millennium*. See p. 1 of this study.

ebooks and ebook publishers. Rather than render ebooks more susceptible to illegal use, open standards promote user investment in ebooks and create a robust ebook market. Ultimately, ebook readers, distributors, and publishers have a similar goal: the development of an accessible, flexible, participant-friendly ebook environment.

APPENDICES

Appendix A: Ebook Format Types

Appendix B: Ebook Distributors

Appendix C: UBC Press Freelancer Guidelines for Collection Point

Appendix A: Ebook Format Types

Platform or Company	Format
Plain text	TXT
Adobe	PDF
FictionBook	FB2
Ebrary	EDF
Gemstar	PRC, IMP, or REB
Hiebook	KML
Web	HTML or XHTML
IDPF	OPS
OpenDocument (graphs, etc.)	ODF or ODT
Microsoft Word	DOC or RTF
MS Reader	EBO or LIT
MobiPocket	MOBI
Palm	PDB or PRC
StarEbook	STK
Sony E-reader	BBeB (also LRF) or RTF

Appendix B: Ebook Distributors⁴⁷

Distributor	Service Type	Formats Accepted from Publishers
UBC Press		
CEL	Library Aggregator	uPDF
Google Book Search	Ebook Vendor	uPDF
Ebrary	Library Aggregator	uPDF
Others		
Amazon Search Inside	Ebook Vendor	uPDF
Amazon Upgrade	Ebook Vendor	uPDF
Baker & Taylor	Library Aggregator	uPDF
BookSurge	POD	Custom
Books 24x7	Specialized IT	Custom
Ebook Library	Library Aggregator	uPDF
ebooks.com	Ebook Vendor	uPDF
Ebsco Booksources	Library Aggregator	uPDF
Follett	Library Aggregator	uPDF
Gale Virtual Reference Library	Library Aggregator	Custom
Knovel	Specialized Scientific	uPDF
Lightning Source	POD	Custom
Microsoft Live Book Search	Ebook Vendor	uPDF
MobiPocket	Ebook Vendor	Custom
MyiLibrary	Library Aggregator	uPDF
NetLibrary	Library Aggregator	uPDF
Numilog	Ebook Vendor	Custom
O'Reilly Safari	Specialized IT	uPDF
Overdrive	Library Aggregator	Custom
Printorium	POD	Custom
ProQuest	Library Aggregator	Custom
Questia	Library Aggregator	uPDF
Sony Connect	Ebook Vendor	Custom
VitalSource	Security Software	Custom
World eBook Library	Library aggregator	uPDF
WOWIO	Ebook Vendor	Custom

⁴⁷ This list of distributors was compiled by the author from information provided in codeMantra, "Ebook Distributors," non-confidential informational PDF file sent to UBC Press, 2007.

Appendix C: UBC Press Freelancer Guidelines for Collection Point⁴⁸

Typesetters

First proof

- upload to CP as "Uncorrected Proof PDF"
- specify email notification to PE

Subsequent rounds of proof:

- upload revised PDF of entire book to CP as "Uncorrected text proof PDF" (this will overwrite previous version of proof that had been uploaded)
- specify email notification to PE

When text is ready to go to the printer: upload only

- high-res PDFs of text and alternate CIP page to CP (as "Text complete PDF for printer hi res" and "CIP pg for paperback PDF")
 - except for the alternate CIP page, make sure that entire text is in a single PDF file (do not separate FM from main text, etc.)
- if there is a plates insert that will print separately, upload this as well, as "Plates insert for printer PDF"
- text information form for the printer
- specify email notification to PE

⁴⁸ UBC Press, "CP Procedures for Freelancers," informational document for freelancers, 2007.

If there are corrections to DBLs

- email replacement page(s) to PE
- integrate corrections into native files; zip all native files and upload to CP as "Text complete [application name]" and "CIP pg for paperback [application name]"
- generate fresh high-res PDF and upload to CP as "Text Complete PDF for printer hi-res" and "CIP pg for paperback PDF" if this changed (replacing original version of each)
- generate low-res PDF of complete book (with hardcover CIP page) and upload to CP as "Text complete PDF final low res" [if there is a plates section or other separate component in the interior, include it in this low-res PDF]
- specify email notification to PE *and* to info@ubcpres.ca

If there are no corrections to DBLs, PE will contact you to request these uploads:

- Low-res PDF "Text complete PDF final low res" (with hardcover CIP page) [if there is a plates section or other separate component in the interior, include it in this low-res PDF]
- Native files, zipped together, as "Text complete [application name]"
- specify email notification to PE *and* to info@ubcpres.ca

Designers

When front-cover design has been approved, upload to CP:

- PDF, as "Cover front PDF"
- full-size, 300-dpi TIFF, as "Cover front TIFF"

Included here with the permission of UBC Press.

- specify email notification to PE *and* to congdon@ubcpress.ca

When files are ready to go to the printer, upload to CP:

- zip together all native files for hardcover and paperback (if applicable)
 - in this zip archive, include low-res reference PDFs of the full cover mechanicals (include "UBC Press Archive" in the file name of these PDFs)
 - also include spine-width templates
 - upload as "Cover files zipped"
- upload Cover info form for printer
- specify email notification to PE

REFERENCE LIST

- Access Copyright. "Incidental Photocopying." *Access Copyright*.
<http://www.accesscopyright.ca/Default.aspx?id=93> (accessed July 20, 2007).
- Adobe. "Products and Services." *Adobe*. http://www.adobe.com/products/#acrobat_family (accessed July 30, 2007).
- American Association of Publishers. *Digital Rights Management for Ebooks: Publisher Requirements*. New York: American Association of Publishers, 2000.
<http://publishers.org/digital/drm.PDF> (accessed July 31, 2007).
- Andersen Consulting. "Reading in the New Millennium: A Bright Future for E-book Publishing." New York: American Association of Publishers, 2000.
Summary available at http://www.accenture.com//rdonlyres/78BBA6F9-7370-49A3-AEFF-46C77CC6669A/0/medi_aap.PDF (accessed July 29, 2007).
- Baker and Taylor, Inc. *Baker and Taylor, Inc.* <http://www.btol.com> (accessed July 19, 2007).
- Brand, Megan. *Outsourcing Academia: How Freelancers Facilitate the Scholarly Publishing Process*. MA Thesis, Simon Fraser University, 2005.
- Book Study Information Group. "ONIX (Online Information Exchange)." *Book Study Information Group*. http://www.bisg.org/onix/onix_faq.html (accessed August 11, 2007).
- . "BISAC Metadata Committee." *Book Study Information Group*. <http://www.bisg.org/bisac/metadata/index.html> (accessed August 11, 2007).
- codeMantra. *CodeMantra*. <http://www.codemantra.com> (accessed July 10 – August 14, 2007).
- . "Ebook Distributors." Informational PDF file sent to UBC Press, 2007.
- Canadian Research Knowledge Network. "Digital Content Infrastructure for the Human and Social Sciences (DCI) Project Implementation Process: Frequently Asked Questions (FAQ)." *CRKN/RCDR Canadian Research*

- Knowledge Network*. <http://researchknowledge.ca/en/programs/documents/DCIConsultationFAQ-English.PDF> (accessed August 2, 2007).
- . "Electronic Content Expansion Project: Request for Proposal." Informational package sent to publishers. Ottawa: CRKN, 2005.
- . "What Is the CRKN?" *CRKN/RCDR Canadian Research Knowledge Network*. <http://researchknowledge.ca/en/index.jsp> (accessed August 1, 2007).
- Davy, Diane. "The Impact of Digitization on the Publishing Industry." Report Prepared for the Association of Canadian Publishers, March 31. Ottawa: Association of Canadian Publishers, 2007.
- Ebrary. *Ebrary*. <http://www.ebrary.com> (accessed July 17- August 1, 2007).
- Fluckinger, Don. "Adobe, AIIM Make ISO Push for PDF." *PDFZone*. <http://www.PDFzone.com/article2/0,1895,2088267,00.asp> (accessed August 1, 2007).
- Gibson Library Connections. "Canadian Electronic Library." *Gibson Library Connections*. <http://www.gibsonlibraryconnections.ca/CPC%20Partners%20FAQ%20Sheet.htm> (accessed July 13, 2007).
- Gibson Publishing Connections. "Introduction to Collection Point for Canadian Publishers." Informational booklet for publishers, 2007.
- International Digital Publication Forum. *International Digital Publication Forum*. <http://www.idpf.org> (accessed July 21-Aug 5, 2007).
- McCoy, Bill. "Adobe Delivers Digital Editions 1.0." *Bill McCoy*. http://blogs.adobe.com/billmccoy/2007/06/adobe_delivers.html (accessed August 2, 2007).
- . "PDF Slated to Become ISO Standard, 'All Inclusive.'" *Bill McCoy*. http://blogs.adobe.com/billmccoy/2007/01/pdf_slated_to_b.html (accessed August 1, 2007).
- . "Steve Jobs: 'Eliminate Music DRM!': So, What About ebooks?" *Bill McCoy*. http://blogs.adobe.com/billmccoy/2007/02/steve_jobs_elim.html (accessed August 1, 2007).
- Mooney, Stephen. "Interoperability: Digital Rights Management and the Emerging Ebook Environment." *D-Lib Magazine* 7. 1 (Jan 2001).

<http://www.dlib.org/dlib/january01/mooney/01mooney.html> (accessed July 31, 2007).

MyiLibrary. *MyiLibrary Ebook Platform*. <http://www.myilibrary.com> (accessed July 19, 2007).

NetLibrary. *NetLibrary*. <http://www.netlibrary.com> (accessed July 15-Aug 1, 2007).

NetLibrary. *PDF Submission Guidelines*. <http://a1835.g.akamai.net/7/1835/276/95d75154a85342/extranet.netlibrary.com/extranet/publisher/specifications/documents/PDFSubGuidelines.PDF> (accessed July 18, 2007).

Simon Fraser University Library. *Library Catalogue*. <http://troy.lib.sfu.ca> (accessed July 22, 2007).