A PExODYSSEY: BIBLIOGRAPHIC DATA MANAGEMENT AND THE IMPLEMENTATION OF PExOD AT THE DUNDURN GROUP

by

Heather Sanderson
Bachelor of Arts, Honours, Queen's University, 2003

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

MASTER OF PUBLISHING

In the
Faculty of Arts & Social Sciences

Publishing Program

SIMON FRASER UNIVERSITY

Fall 2004

Approval

Name: Heather Sanderson

Degree: Master of Publishing

Title of Project: **A PExODYSSEY: BIBLIOGRAPHIC DATA MANAGEMENT AND THE IMPLEMENTATION OF PExOD AT THE DUNDURN GROUP**

Supervisory Committee:

Dr. Rowland Lorimer
Senior Supervisor
Publishing Program

John Maxwell
Supervisor
Publishing Program

Kirk Howard
Publisher
The Dundurn Group

Date Approved: December 07/04

SIMON FRASER UNIVERSITY

PARTIAL COPYRIGHT LICENCE

Abstract

In the summer of 2004, The Dundurn Group, implemented a new database technology called the Publisher's Extensible ONIX Database (PExOD). PExOD is a bibliographic database developed at the Canadian Centre for Studies in Publishing in 2002. The decision to adopt this technology at The Dundurn Group came from the increasing digitization of the book publishing industry, the use of ONIX for exchanging bibliographic data with trading partners, and the lack of any formal system for dealing with data at the company. This report describes the implementation, integration, and impact of PExOD on the publishing processes at The Dundurn Group, successes and problems of implementing the software, and suggested changes to PExOD, while also explaining the context of PExOD within the structure and problems inherent to the supply chain in Canada.

Dedication

To Mom, Dad, and Kevin for your constant love and support.

Acknowledgements

I would like to thank Dr. Rowland Lorimer and Kirk Howard for seeing me through this project from day one. John Maxwell for his continued effort to answer hundreds of my emails and questions; for the meetings, phone conversations, and brainstorming sessions that helped me complete this project. To Beth Bruder for prompting me towards this topic and for providing a solid learning environment in which to work. A special thanks to Jennifer Easter and Anne Choi, without whom I surely would have lost my sanity. To Ian Payton for answering many questions, and to the rest of the staff at Dundurn for being so welcoming and willing to go on the adventure of this PExODyssey together.

Table of Contents

List of Figures and Tables

Introduction

In this report I explore the interrelationship between bibliographic data management, the publishing process, and the use of database technology at The Dundurn Group. I focus on how bibliographic data and author information has been managed at the company in the past, and how a new database system (PExOD) transformed both the existing processes within the publishing house as well as its interaction with partners in the publishing industry. Though Dundurn had been keen on the idea of using PExOD since its inception, there are several reasons I indicate as to why they were not able to implement the system prior to the summer of 2004.

The information in this document was collected between May and December 2004 from publications, articles, websites, emails received by and sent from Dundurn employees, a questionnaire that was administered to employees, and the results of a period of implementation, trial, and error working with PExOD. To provide context to my argument I offer additional information regarding the history of bibliographic data, ONIX, and PExOD within the overall framework of Canadian publishing. I also examine a theoretical application of password-protected departmental sections within the PExOD software. Later in the report I focus on the impact that PExOD has had at Dundurn, and the plethora of possibilities that it possesses for the company from driving the website to increased sales through e-commerce.

Bibliographic Data Management Background and the Creation of ONIX

What is Bibliographic Data?

The term "bibliographic data" refers to product information about a book that has been published in the past, or will be published in the future. This information is commonly separated into core data and rich data. Core data include information such as the title of the book, ISBN, and the author's name whereas rich data are data that expand upon the essence of the book, such as copyright information, cover image, author biography, book description, page count, dimensions, and format of a particular book. Essentially, bibliographic data is a long list of the qualifying and unique information that indicates everything anyone would want or need to know about a book. These metadata (i.e., data about data) are useful, first and foremost, because they distinguish one item from the next.

Title information about different books has always been at hand in the publishing industry, but it seems to have been overlooked as something that could be a useful tool for publishers in various capacities. As a result, structural and systematic ways of dealing with or organizing bibliographic data have not always been (and still are not always) a priority in publishing companies. For example, a publisher may tell a reviewer about a book and its author without realizing that this information is something that should be categorized and kept on record: the information is common knowledge to the publisher,

but may not be shared adequately with members of the supply chain. Such interactions are common and ultimately, they can create a jumble of incomplete data in a publishing company. It would be extremely advantageous for this information to be preserved, archived, and organized in such a way that the entire staff and other trading partners have ready access to product information: if valid data are not accessible they are of no value to the company.

Bibliographic data are generated at a publishing house, and can be used for multiple and varying purposes. But for them to be used, they must be valid, organized, and easily accessible within the publishing house. Availability of and access to valid data ensures that the entire staff is knowledgeable and has up-to-date information (often a problem, as it is constantly changing) for every title as it is created and after publication. A good data management program for publishers goes beyond merely storing basic data about a book; it should also have the capacity to store cover images, author pictures, review quotes, and contact information for authors and clients. Were such a system in place, a publicist no longer would have to find a file folder with hard copies of past reviews, or ask a designer for images needed to complete a press release. These items can be stored in a logical place in a database. Implementing a data management software program increases efficiency in the completion of daily tasks such as creating press releases. Similarly, the amount of time put into large projects such as the creation of a catalogue also diminishes when a publishing firm begins to use a data management tool. But ready accessibility to accurate bibliographic information is probably most useful as a sales tool.

Product information is the base of publishers' catalogues — the primary tools used to help sell a list of books in a particular season. Product information is also printed on the dust jacket of a book to convey information about the product to customers (with the goal of persuading them to purchase a book). Essentially, product information is used (in various formats) to promote a book to retailers and to the end consumer. In short, although bibliographic data are important to securing sales for the publisher, they are intrinsic to all sectors and levels of the publishing industry.

Bibliographic Data and The Supply Chain

Each trading partner serves a specialized function within the supply chain and, subsequently, in the flow of information. Merely obtaining data is the beginning of the process. Once obtained, trading partners put the same data to work in slightly different ways. Everyone involved benefits, from increased efficiency in their jobs to augmented sales. Since business partners use overlapping sets of data the accuracy of these data is essential if a publishing house is to succeed. Moreover, while this information could once change hands via a paper hard copy, fax, or over the phone, the sheer volume of titles that are currently in print (113,400 between 2000-2001 in Canada)[1], paired with the availability of computerized database technology, make the creation of readily communicated digital data files both a reality and a necessity.

The structure of the supply chain in publishing necessitates and relies upon the production, organization, and maintenance of accurate data to help ensure a steady flow of information through the intricate web of parties involved: publishers and their staff, distributors, wholesalers, retailers, and consumers. Each of these entities fulfils a specific

function and has slightly different demands on the needed data. For example, publishers often use distributors to fulfil orders to retailers. These distributors require accurate product information from the publisher, which is then managed at the warehouse to ensure that the correct books are in stock and orders can be fulfilled accurately and efficiently. The distributor must also record customer information, as well as how many units of each book are shipped and returned, in order to generate reports that provide the publisher with profit and loss information to track sales, and information upon which royalties can be based. If a distributor does not have accurate title information, it cannot process any orders, and no sales take place. It is the publisher's responsibility to ensure that the distributors have accurate data for all titles that it distributes. For these reasons, it is important for both the distributor and publisher to use the same qualifying bibliographic data to ensure that information can be generated and communicated clearly.

Retailers and wholesalers also play a crucial role in the Canadian supply chain. The retailer component of the supply chain in Canada is dominated by one chain store: Chapters, whose business formula depends on the ready flow of bibliographic information. Its impact on the industry has been such that Roy MacSkimming states "Chapters Inc. has transformed the Canadian retail book market so radically that it's startling to recall the corporation has existed only since 1994" (MacSkimming, 2003: 360). Chapters sprung out of an amalgamation of the bookstores Coles and SmithBooks, and later the Toronto-based Indigo Books and Music purchased the company (MacSkimming, 2003: 361). As a chain superstore, Chapters is able to draw higher discounts from publishers, and therefore offer lower prices to customers. The presence of Chapters in the market and its aggressive competitive policies — for example, building

near established independent stores — has resulted in the closures of many independent booksellers.

Major wholesalers such as National Library Services also need accurate bibliographic data. They purchase books at a high discount from publishers for their clients, and then distribute these books themselves (to school libraries across Canada). Wholesalers and retailers require bibliographic data so that they can determine the specifics of what they are ordering from a distributor or publisher (for example a trade paperback edition of a book versus a hardcover copy), and to ensure they have access to the information required to aid in sales. Bibliographic data are also used to record sales and returns information, and to help promote a book either online or through other advertising and promotions. Traditionally this information has been typically obtained through publisher's catalogues, by meeting with sales representatives, and by consulting cumbersome publications like *Canadian Books in Print*. Marketing and publicity campaigns function as reminders to reinforce the existence (and saleability) of particular titles. As these basic data pass from a publisher to other members of the supply chain, it is imperative that they be complete and accurate. One hiccup in the data can result in bookstores not being able to place orders, as happened at Dundurn with a book called *Treasure and Intrigue*. The incorrect ISBN was sent to the retailers, which resulted in no orders for the book from Chapters until the mistake was rectified months later.

Online booksellers have created a retail environment that depends completely on the digital bibliographic data generated by publishers. The first major company to create a model of electronic book selling was Amazon.com (in 1995), but others soon followed suit. For example, to capitalize on this new source of revenue, Chapters launched an

online store in 1998 (MacSkimming, 2003: 361). Many publishers have also started to use their websites as a means by which they can generate revenue through direct online sales to customers as well as a marketing tool. This form of retail was particularly interesting to the publishing industry, because online shoppers could no longer pick up a book in a store and read the jacket copy; they had to find the information on a website instead. As online booksellers started to become a recognized source of income for publishers, and as larger chain retail stores began to emerge, there was a transformation not only in the way the book selling business was conducted, but also in the flow of information: its significance took on new importance.

The final link to the supply chain is the main reason why the rest of the intricate web is in place: the consumer. Consumers need access to bibliographic information in order to make an educated decision about what they wish to purchase. The more information a customer has about a book, the more likely it is that he or she will purchase it. With sixty per cent of book purchasing decisions made in the store, the information found printed on a book cover — jacket copy, author picture, quoted review extracts, content summary — contributes to the final decision to buy a book (Barnes and Lorimer, 1997: 17). For online book purchases, this is even more the case.

Technology, the Supply Chain, and ONIX

Technology started to play a pivotal role in sales and bibliographic data management when the personal computer was first introduced to the office in the 1980s, and many publishers began to produce spreadsheets containing title information in programs like Lotus 123 and later Microsoft Excel to categorize and manage their data.

In the 1990s, the Internet started to bring new possibilities such as online book sales, and new means of advertising and publicity to book publishers. As online retailers began to emerge, a new way of purchasing books was realized.

As computers, the Internet, and advanced databases created an environment of digitization and automation in the process of book publishing, the interaction and interrelationship between the publishers, distributors, wholesalers, retailers, and consumers began evolving (and is still in a state of transition). Through electronic communications technology, information could be transferred at a rapid pace: exchanged easily, instantaneously, and internationally. For example, books can now be purchased on-line from anywhere in the world and sent directly from a warehouse to a customer's home. This process came about wholly because of the Internet and it created, or rather, entrepreneurs used the Internet to create a retail environment that bypasses the traditional distributor and bookseller. This development, in turn, encouraged digital automation and an expanded global supply chain. It also brought forward the question: how can publishers get product information to the supply chain in an efficient and timely manner?

The flow of product information in digital form within the supply chain, and the fact that books can be purchased online are new practices that force publishers and customers to adapt to a new retail environment and new processes by which bibliographic information is created, managed, and distributed. To deal with this digital revolution companies found ways to process data differently. The result was that most companies developed completely different ways of storing bibliographic data in their own databases, or they created qualifiers for data that were important for their own purposes. However, these qualifiers did not necessarily match with those of their trading partners.

The major problem that arose from working within a supply chain that had no formalized structure or body of governing rules was that each interaction between any two parties could require information to be formatted differently, or often required an entirely different set of qualifying information to exchange hands. A certain wholesaler, for example, could ask that each title be entered in a spreadsheet under specific headings such as title, subtitle, and author name, whereas another retailer may have developed a slightly different or more complex system of coding for this same information. These inconsistencies, common within the industry, are time-consuming for the publisher, as publishers often have to fill out different templates, forms, or spreadsheets and then transmit this information to each client: the communication process is slowed or, in some cases, halted as a result. It is also enormously inefficient. With the constant flow of information occurring in the chaotic and expanding structure of the supply chain, it soon became evident that processes and standards needed to be developed to ensure that all parties could obtain basic bibliographic data without needless duplication of effort.

The fundamental solution to this problem is data standardization, and the standard protocol being adopted by the book industry is called ONline Information EXchange (ONIX). It was first released was in January 2000. According to the Book Industry Study Group, ONIX is "a standard format that publishers can use to distribute electronic information about their books to wholesalers, e-tail and re-tail booksellers, other publishers, and anyone else involved in the sale of books" (BISG, 2004). The overall goal of ONIX is to "standardize the transmission of product information so that wholesalers, retailers and others in the supply chain will all be able to accept information that is electronically transferred in ONIX International format" (BISG, 2004). ONIX is

essentially a building block that offers the ability to standardize all data which, as a result, improves efficiency.

The development of ONIX arose out of both the alterable problems regarding the interruption of data flow and the inaccuracy of information provided within the supply chain and the advancement of technology in the late 1990s and early 2000s. In 1999, the American Association of Publishers (AAP) "worked together with the major wholesalers, online retailers, and book information services to create a universal, international format in which all publishers, regardless of their size, could exchange information about books" (EDItEUR, 2004). This initial study conducted by the AAP led to the realization that standardization of bibliographic information was needed and would greatly benefit the industry globally. As a result, the AAP released ONIX 1.0 in January 2000. From this point, the ONIX standard was then "developed and maintained by EDItEUR jointly with Book Industry Communication (UK) and the Book Industry Study Group (US), and with user groups in Australia, Canada, France, Germany and the Republic of Korea" (EDItEUR, 2004). The fact that so many countries are involved in the development of ONIX means that it has the potential to become the overarching, all-encompassing, world-wide standard.

Currently there are two standards for ONIX: Level 1 requires the most basic data, and is in place to ensure that small publishers can exist within the supply chain without an extremely sophisticated data management system; Level 2 allows the publisher to input more extensive data elements (EDItEUR, 2004). ONIX allows for over two hundred data elements, "each of which has a standard definition, so that everyone can be sure they're referring to the same thing. Some of these data elements, such as ISBN,

author name, and title, are required; others, such as book reviews and cover image remain optional" (EDItEUR, 2004). Though the main overall goal remains the same, each country with a vested interest in ONIX may govern its use differently, especially with regard to which material is mandatory and optional, to ensure that this technology best meets the needs of the supply chain within which they exist.

In Canada, the use of ONIX is stewarded by BookNet Canada. BookNet Canada is a "not-for-profit organization dedicated to innovation in the Canadian book industry supply chain. It establishes Canadian standards to improve bibliographic databases, distribution and inventory management" (BookNet, 2004). They have developed three different standards for bibliographic data in Canadian publishing: gold, silver, and bronze. The bronze level is the most basic level of data. With twenty-six basic elements (between nine and nineteen of which are required by major retailers) that can be sent in a spreadsheet, all publishers are required to meet this level of information in order to trade in the supply chain. Some publishers operate at the silver level, which includes the same elements as the bronze, but is transmitted through the ONIX 2.1 standard (EDItEUR launched ONIX 2.0, providing the use of full XML into the system, in 2001 and another upgraded version, ONIX 2.1, was released in 2003) (Sasmor, 2004: 25). The gold level includes the prior two levels plus addition information for marketing purposes such as cover images, page count, reviews, or a description of the book.

ONIX standards are constantly being updated and are created by EDItEUR. EDItEUR is "an international organization with European origins that coordinates the development, promotion, and implementation of EDI (Electronic Data Interchange) in the books and serials sectors. It helps maintain ONIX for European and international

communities" (EDItEUR, 2004). All countries involved in the utilization of ONIX communicate with one another, and with EDItEUR, to improve ONIX the system: to ensure that ONIX is being employed properly, and is used to the advantage of each country.

Although the use of ONIX is logical in theory, it is a difficult system to implement at all levels of the industry. As it was produced in 2000, and the time of this report is 2004, it is still a relatively new system. With any new system, aspects are still being improved upon and altered in order to ensure that its potential is maximized. Furthermore, this change has ultimately affected all publishers, as they must find a way to ensure that their bibliographic data are formatted properly and adhere to a newly developing, and evolving, world-wide standard.

From ONIX to PExOD

How ONIX works

In order for ONIX to work effectively, a publisher must first ensure he or she stores, formats, and manages data digitally. He or she must then transfer these data in an ONIX message. ONIX governs how these data are formatted and exchanged between computers. This message is defined by a series of tags (a different tag for each element within ONIX), is written in XML (eXtensible Markup Language), and conforms to a specific set of rules or Document Type Definition (DTD)(EDItEUR, 2004).

XML is a markup language that goes beyond the capabilities of HyperText Markup Language (HTML) with regards to creating a structure for content. XML "offers a unique combination of flexibility, simplicity, and readability by both humans and machines" (St. Laurent, 1999: xxiii). Essentially, "XML offers software developers the opportunity to create documents with built-in frameworks that make it much easier to create consistent results time after time, capable of carrying useful, and reusable data" (St. Laurent, 1999: 53). XML is considered to be the evolving language for data exchange (Lawrence, 2004).

ONIX messages are generated in XML because it "is optimized for creating complex documents and transmitting and exchanging data between computers." XML is also easily readable by both humans and computers (EDItEUR, 2004). To export the data in ONIX, they must first be collected and stored in a database, and then organized within

the structure that an XML message creates. Once the message is generated, it is transmitted across the Internet via email or a File Transfer Protocol (FTP). A single message can transfer information about several books at one time.

While the principle of ONIX and a universal standard was something that the supply chain required in order to improve the publishing business, the transmittal of an ONIX message in XML as a mandatory method is still forthcoming. Retailers and wholesalers typically have adopted ONIX field names as a means by which information is standardized, but still require different forms or templates in Microsoft Word, Access, or Excel to be filled out and transmitted via email, or filled out in online forms. In order to continue the evolution and implementation of ONIX, and to increase productivity and efficiency within publishing houses, more retailers need to accept ONIX messages in XML. Publishers need to adopt technology capable of producing such a message. Based on the ever-increasing adoption of automation and technology in the Canadian supply chain, it seems that most of the major booksellers who have made ONIX a required standard are now accepting XML feeds. It seems likely that they will make this a mandatory process for accepting information in the near future.

ONIX and Retailers

From the inception of ONIX in 2000, the process of data exchange became highly accelerated and it was not long before many members of the supply chain (mostly retailers) started to support this format of data exchange. With major retailers such as Chapters, Amazon, Ingram Book Company, Barnes & Noble, and R.R. Bowker requiring that all data from publishers be submitted to them in ONIX between 2002 and 2004,

publishers had to find a way to incorporate this system into their data management and sales processes.

Orders that were once accepted via fax or over the phone became a thing of the past as computers began a digital revolution within the supply chain. For instance, Chapters now insists that publishers supply data in the ONIX standard months before a book goes to print. As Chapters has "market shares of between 70% and 83% in the four major cities [Vancouver, Calgary, Toronto, and Montreal], Chapters Inc. clearly dominates the book retail market and it has the power to dictate change in the industry" (Evans, 2001: 11). This means that if publishers want to maintain a relationship with this larger retailer, they must adhere to certain ordering and data sharing requirements. This means that publishers had to find a way to manage the data internally, digitally, and inexpensively, which led to various solutions including the development of the Publisher's Extensible ONIX Database (PExOD).

What is PExOD?

PExOD is an online, ONIX-compliant, database system developed at the Canadian Centre for Studies in Publishing (CCSP) at Simon Fraser University (SFU) funded by the Department of Canadian Heritage (DCH) as part of the government-based Supply Chain Initiative. The CCSP recognized that Canadian publishers would require a means by which they could organize and export data in ONIX. As no such databases existed in the early 2000s there was a gap that could be filled through research and development of database software that co-existed with newly emerging standards.

The CCSP started developing PExOD in 2002, and has worked to improve the software's functionality and usability since its inception with three major upgrades. The target user for PExOD, in the eyes of the CCSP, is small to mid-sized publisher. As of September Don Smardon, the Webmaster of New Society Publishers, uses PExOD to manage all of the company's bibliographic data. Brenda Martin, Heritage House Publishing's Marketing Manager, and Andrew Wooldridge, Orca Book Publishers' Editorial and Sales Manager, are currently using PExOD in some way. New Society has developed the software to the point that it is used as the database that drives its website.

As PExOD is web-based (meaning that users can access the database online if they have the proper user name and password as opposed to a database that is stand-alone) it is accessible around the world on different kinds of computers, can be used in any place at any time, and is capable of running on an individual computer, an in-house LAN, or a third-party host such as an ISP. While PExOD itself is not distributed under an open-source license, it is built on top of open-source technologies (PHP and MySQL), and as such the "source code for the software [is] available along with the object code" (Lessig, 1999: 103). Open-source technology is based on the idea that "others [can] explore how a program [is] implemented and learn from that example how better to implement the protocol in the future" (Lessig, 1999: 103). This means that PExOD is not proprietary software and, as such, anyone that uses the program on their own server can access the structure and code of the program in order to improve or build upon what is already there. Another advantage of PExOD is that it can run on different platforms (such as UNIX and Windows), so it can work with existing systems in a publishing house.

Moreover, it is not as costly as other proprietary database software that is currently on the market.

Functions of PExOD

PExOD consists of two main interfaces with different functions: administrative and public. The public interface allows the user to search for titles, an index, and a browse-by-subject option. Once a search has been generated, and a title selected from the search results, a template of a web page appears with some of the pertinent information about the title. A publisher can choose how her title information appears on the public by customizing a set of PHP-based web templates.

The administrative side of PExOD allows the user to enter title information by either directly inputting it into the correct field, or by uploading a spreadsheet of data. The administrative interface contains over 200 fields for information that is ONIX-compliant in that the fields where information is entered are tagged with the corresponding ONIX codes. This method of tagging is done in such a way that PExOD can create an ONIX message in XML based on the information that has been entered and this information can be exported in the ONIX standard. A publisher can enter or update data at any time, and changes are automatically manifested on PExOD's public interface. Beyond the basic function of storing bibliographic data, publishers can store and manage author and other contact information, add rich data such as cover images, author photos, reviews, and book descriptions. Once data are entered and validated, they can be exported directly as an ONIX message in XML to any online bookseller or retailer who is set up to accept this message. In July 2004, PExOD was updated to include the new ONIX 2.1

upgrade, meaning that a publisher can export data in ONIX 1.0, 2.0, or 2.1. Though most retailers are not yet accepting the newest 2.1 feed, they will accept the prior versions of ONIX.

Although the initial principle for PExOD was to act as a bibliographic management tool for small to mid-sized publishers, the software can be implemented to fill other voids in a publishing company. As the program is web-based, all Internet-connected employees can access the database. This means that PExOD can also act as a central server to store information in a company where there is no designated server or network. A central server for information means that employees do not have to spend time searching for data time and time again. PExOD is also able to act as the main database for a company's website. For some smaller publishers, this means that they can create a website using the templates in PExOD where they may not have had the resources to build a website in the past. See Figure 1 for a list of the basic functions and advantages of PExOD.

Figure 1: Functions and Advantages of PExOD

- Uses the ONIX standard
  - Import a spreadsheet from Excel or Advance with title information
  - Searches database by title, author, title info pages
  - Shows a complete title list, contributor list, enterprise list, account list, contact list and custom category list.
  - Add a new title not in a spreadsheet.
  - Includes marketing and promotional information about a title.
  - Exports information to be used by booksellers such as Amazon and Chapters.
  - Exports information to distributors.
  - Manages author information, including photos and biographies.
  - Designates specific imprint for a title.
  - Links related products.
  - Adds and manages multiple supplier and price details (in all currencies).
  - Use as database for company website.
  - Acts as a central information server with accurate information about all current and backlist titles.
  - Reduces the duplication of tasks and tracking down certain computer files.
  - Ensures that the information is in accordance with ONIX standards.

PExOD and the Supply Chain Initiative

To understand the role of PExOD in Canadian publishing, it is necessary to understand the context of the Supply Chain Initiative. Developed by the Department of Canadian Heritage as part of the Book Publishing Industry Development Program (BPIDP),[2] the Supply Chain Initiative "aimed to support the adoption of new technologies making physical distribution more efficient and helping publishers and booksellers to manage the inventory that constantly flows between them" (MacSkimming, 2003: 387). The Supply Chain Initiative was developed in 2001 as a

means to solve structural weaknesses inherent to the supply chain. The major supply chain problems were the unmanaged formatting of title information, as discussed in section one, and poorly managed distribution of this information: both of which led to missed sales opportunities. The Supply Chain Initiative has three main objectives:

1. improving the accuracy and comprehensiveness of bibliographic data on books for sale in the Canadian market,
2. promoting standards-driven electronic document interchange, and
3. promoting access to accurate and timely data on book sales in the Canadian market (DCH, 2004).

These measures, it was hoped, "would cut down on the waste and expense of returns, thus making the industry more profitable" (MacSkimming, 2003: 387). To accomplish these goals, DCH provides funding to book publishers, retailers, distributors, as well as supporting the creation of special projects that will aid the supply chain in some way. PExOD emerged as part of the solution to these problems and is funded, in part, by DCH.

PExOD provides publishers with a system that "is designed to allow book publishers to create a bibliographic database of their own titles, and titles they handle, from which data elements can be easily retrieved" (Lorimer et al, 2004: 2). PExOD provides publishers with an opportunity to organize and centralize all of their data in such a way that all employees can access it and ensures that it can be exported to other members of the supply chain in ONIX.

The Dundurn Group, Bibliographic Data Management, and PExOD

Dundurn Press was co-founded in 1973 by Kirk Howard and Ian Low. Since its inception, the press has acquired or created four other imprints: Hounslow Press, Simon & Pierre Publishing, Boardwalk Books, and Castle Street Mysteries. As each of these imprints specializes in different genres, their amalgamation under the umbrella of The Dundurn Group fostered one press with a diverse list. This also meant integrating bibliographic data and author information from various sources.

Creation and Storage of Information at Dundurn

Dundurn is a medium-sized publishing house that has a frontlist of around fifty titles each year. Typically there have been fewer than ten employees until the recent creation of new jobs, which means that Dundurn had a staff of twelve in the fall of 2004. Though it is a mid-sized company, each department (editorial, publicity, sales and marketing, and design) is highly segregated and communication does not always flow between departments. This means that the process of creation and storage of bibliographic data is also highly segregated and compartmentalized. For example, the editorial director has control over the ISBN allocation for new titles, and is the first to acquire the new title and author information. He may enter this information into a spreadsheet, and store it on his computer, but does not distribute it to any other staff member unless they ask him for it. The same problem occurs for obtaining cover images

from the designer or marketing and publicity campaign information: there is no formal mechanism in place to ensure that information is circulated or shared with other employees. Unless an individual needs to find something to complete a task as simple as returning a phone call to a client or author to more complex projects such as creating the catalogue, the information often stays with the person with whom it originated.

Some attempts have been made to ensure that information remains intact, and is accessible, but processes are never forced upon the company as a whole. As a result, different systems are used by different employees, resulting in gaps in the data, confusion, and inefficiencies. Author contact information is a good example of this. Contact information is often only stored on the original hard copy of a contract, which can be found in large filing cabinets in the office; in incomplete and outdated publicity binders; or in Kirk Howard's Outlook Express address book. If an address changes, these systems aren't always updated, but one employee may possess a post-it note with the information scribbled on it. Each of these different approaches to storing contact information have been started by different staff members at different points in time; none have been maintained, and none of them are complete.

There has never been one overarching, holistic approach to data organization at Dundurn. The lack of an organized, systemic way of dealing with data at Dundurn creates inefficiencies in completing tasks, as well as causing unnecessary havoc and frustration. Moreover, one can imagine the implications and problems this system would cause if a staff member decides to leave the company: all of the information that he or she was once responsible for becomes is potentially lost to the rest of the company.

Production Information Exchange at Dundurn

Weekly production meetings at Dundurn bring the entire staff together to update each other on where a new title is in the production process. These meetings also provide a venue for each person to ask another department for the information that he or she needs to move forward with projects, such as whether a cover image is finalized, or if the page count or publication date is to change. Traditionally, after these meetings, everyone returns to their desks and begins to email the requested files to different employees.

The sales staff are responsible for exporting all of the title information data such as ISBN, title, author name, price information, and the description of the book to all of Dundurn's trading partners. This requires formatting several different spreadsheets in different computer programs, and constantly searching the online retailers' websites to ensure that any title information has been posted and that cover images have been sent and posted to the correct title profiles. Formatting the information in one spreadsheet alone could take at least an hour, not to mention all of the time it takes to hunt down title information and cover images from different employees each time the information is required. As each spreadsheet for each different trading partner contains different headings, and requires information to be formatted a specific way (for example, an ISBN with or without dashes, or the author information with surname first, or given name first), filling out these forms takes hours to complete.

Some trading partners do not require a spreadsheet, but ask for detailed web pages to be filled out, and for the information to be uploaded online, which often takes hours to complete. R.R. Bowker, for example, forces the user through eight different web pages to enter each title. Even though there are many different forms, requirements, and processes

in place, the same information is copied over and over and sent to different companies. Copying the data is a process that is redundant and unnecessary: this process could become more organized through the use of a database and ONIX exports, thus making Dundurn a more efficient work environment.

Dundurn, Data Flow, and the Supply Chain

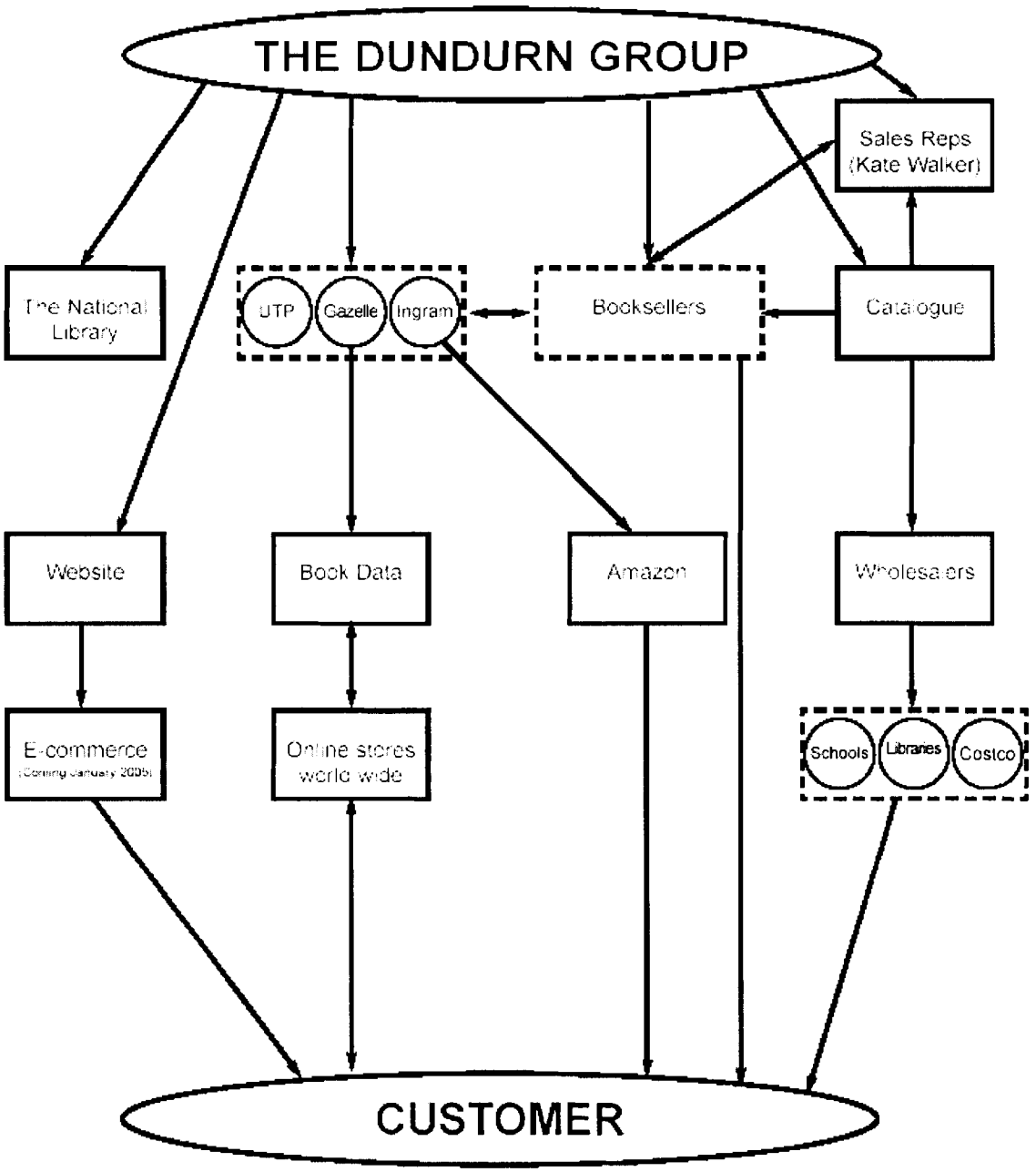
Dundurn's role in the supply chain is one that is quite typical of publishers in Canada. The company maintains strong relationships with independent booksellers, and generates roughly two per cent of their sales from online bookstores, forty per cent from Chapters, twenty per cent from independent booksellers, and twenty per cent from wholesalers (dominantly library wholesalers). The remaining eighteen per cent of sales comes from a combination of special marketing campaigns for different titles, sales to book clubs, author purchases, and special sales. The firm also has three different distributors: University of Toronto Press (UTP) in Canada, Gazelle in the United Kingdom, and Ingram in the United States. Though Gazelle and Ingram accept ONIX messages in XML, they also have alternative means of receiving data from the companies who cannot send ONIX messages: Dundurn is one of these companies. This being the case, each distributor requires Dundurn to transfer the same data in different formats. UTP requires a specific spreadsheet to be completed and emailed to them (though they are aiming to accept ONIX feeds in spring 2005). Ingram requires the information to be manually entered and uploaded on their website (Ingram also feeds information to Amazon.com and other American online retailers). Gazelle has created a database in Microsoft Access that needs to be filled out using American prices, despite the fact that

they are a distributor for the United Kingdom. Gazelle also forwards this information to the data aggregator, Neilsen BookData. BookData creates and maintains "an integrated database of over 5 million English-language titles published around the world (including UK, USA, Europe, Australia, New Zealand and Southern Africa)" (BookData, 2004). Figure 2 illustrates the supply chain processes and flow of information from The Dundurn Group.

Dundurn also receives data referencing the sales and returns of its titles from its distributors. University of Toronto Press (UTP) handles the bulk of Dundurn's distribution and fulfilment of orders. UTP generates a daily report which stipulates the number of units that were shipped and returned, the costs, net sales, gross sales, and a record of which retailer or wholesaler ordered which title. This information is downloaded at Dundurn every morning from an FTP site, and is imported into a program called Client Publishing Electronic Sales System (CPRESS).

CPRESS is a DOS-based program that stores all sales information for current and past titles in each fiscal year end. Some of this information can be downloaded into spreadsheets and then manipulated to create sales reports and forecasts. The final sales information each month is not sent digitally, but is sent on a cumbersome, paper-printed spreadsheet and the total net sales and returns information is then entered manually into yet another software program (Quickbooks) to generate a dummy invoice.

Figure 2: Flow of Data at the Dundurn Group



Dealing With Data: Introducing PExOD to Dundurn

As ONIX started to appear as though it would become the norm for publishers to organize data by, the staff at Dundurn was eager to accept this new format. They saw that there was potential, through standardization and computer-generated automation, to increase efficiency and accuracy when dealing with information and sales. A visit in February 2002 to Baker and Taylor in the United States made them realize that most US companies were ready to accept spreadsheets and ONIX to supply bibliographic data. This realization, paired with the growing influence of Amazon, led them to realize that the flow of electronic bibliographic data ensured that the same information could reach all corners of the world, meaning that anyone could find information about a title, which might result in a sale of that title.

PExOD entered Dundurn's discourse almost simultaneously with the above-described realization. The CCSP's John Maxwell approached the company in April 2002 and offered a demonstration of the software. The culmination of emerging online bookstores, the acceptance of ONIX, and an increase in title production indicated that a solution for bibliographic data management was a necessity at Dundurn. The company was provided with a PExOD demonstration website and access codes to the free software so that they could enter any title information into PExOD 1.0. John also added the company to his email list to ensure that they were kept up-to-date on any new development, upgrades, or frequently asked questions that may arise as PExOD continued to evolve and meet the required needs of publishers.

Though they expressed an interest early on, it wasn't until the summer of 2004 that PExOD started to be implemented at Dundurn. The basic reason for this was that

with such a small staff responsible for so many titles, there was no one who could take time from their hectic schedule to learn how to use new software. Moreover, no one had formal training with PExOD, and could not figure out the complex system at first glance. This made the task seem onerous and easy to put aside to accomplish at some point in the future. A final reason was that though Dundurn's partners in the supply chain were increasingly demanding ONIX-compliant data, they did not require it to be formatted in an XML message. Since the system of creating spreadsheets, though time-consuming, was something that could be completed and was acceptable, it was the easiest option for Dundurn to use when generating and exporting data to its supply chain partners.

Exporting Data via PExOD

A key attribute of PExOD that will further enhance the efficiency of data management at Dundurn is the ability to export ONIX feeds or flat files to other members of the supply chain. Currently there are profiles set up for Amazon and Chapters in PExOD, which allow the user to select as many titles as they need to export and send them to these companies. Profiles can also be created to send information to Gazelle, University of Toronto Press, Ingram, R.R. Bowker, Book Data, and any other firm that accepts ONIX. To create a profile a user must fill in the "Detail" section of the export. This asks for information such as the email address of the recipient, who the message is transmitting from and to, and what type of message it is (for example, an early notification or an update). This ability to export information is a helpful tool as decisions for a book often change right up to the point of publication; most commonly the price, and cover of the book. As the data are often sent to retailers early on, it is difficult to

remember to send updated information, and it is time-consuming to check the websites such as Amazon and Book Data to ensure that the changes have been implemented. PExOD provides Dundurn with the ability to send data extremely early on in the publishing process, and then to update title information in batches as it changes.

The onerous and time-consuming task of filling out spreadsheets and forms can become a thing of the past, replaced by a process that takes all of five minutes to complete. Once the profiles have been created in order to export, you simply have to select the company you wish to export to, and the titles you wish to export by clicking in the boxes next to the title. Once this is done, you click on "Send Selected Titles," and the information is formatted either in ONIX, or a flat file, and is sent via FTP, or email.

Exporting data beyond tests to myself have not taken place yet. The fall 2004 data were not ready in time to be sent, and there is some hesitation on the part of the Dundurn staff to try this new technology. Furthermore, though many of the larger companies are ready to accept ONIX messages, many others that Dundurn works with are not. University of Toronto Press and Gazelle both do not have this capacity, though both are working towards it for Spring 2005. Another key issue that will develop is that publishers now will have to stay updated on which supplier or distributor is now accepting the newest ONIX feed. By this, I mean that some suppliers may accept ONIX 2.1, and others may still require ONIX 1.0. The development of new versions of ONIX still means that Dundurn will have to find out which company accepts which version, or when they will start to accept ONIX. Though this may not be time-consuming, it is still a step that should be taken into consideration.

Website Woes and PExOD

Another important area where a bibliographic database is useful is in website maintenance. Dundurn copies catalogue information each season to their website. The website was created in 1996 because Dundurn recognized that this was an important tool to have to provide title information to customers and trading partners. Dundurn's site holds information for about 400 out of the 1200 books that have been published by The Dundurn Group. Most of the titles are from the years after creation of the website, with selected backlist titles. Each title has its own webpage with pertinent selling features, such as the description, price, cover image, and author biography.

Although this information is available to the web-savvy world, the website has several structural and usability problems. One problem with the website is that it is set up with a flat database file, which makes it difficult to generate searches, the user must know the specific title in order to find it. Another problem is that because different employees have entered information from season to season, the data are incomplete, inconsistent, or absent from the profile for a title. Furthermore, because of the way that the website is set up, when anything needs to be added, changed, or edited, the staff must use Macromedia Contribute to implement changes. Macromedia Contribute is a web-editing program that allows a user to make content and design changes to a web page. This task is typically performed once and, once completed, is forgotten, meaning that the information posted to the website remains static, even though changes to the cover images, price information, or other central pieces of data may change. Moreover, it is a time-consuming process that is only forcing the same data to, once again, be repeatedly drawn upon and supplied to yet another centre.

Currently the website at Dundurn is not in a great state. The company has been working towards a goal with IT personnel to get the website to a point where the design is aesthetically pleasing and the content is current and accurate. These are difficult tasks to accomplish, as the processes and guidelines are not clearly stated. The sales and marketing co-ordinator, Jennifer Easter, and I are responsible for updating the content on the website. This means that whenever the publicist gets information about an event, she forwards an email to Jennifer and myself who, in turn, enter this information on-line using Macromedia Contribute. The IT staff are responsible for creating the overall design of the website, but this design has become outdated in terms of the information that it presents to the public, and in its design. As there is only one person working on the website design and he gives one day a week (30 hours a month) to Dundurn, the company simply does not dedicate sufficient resources to the maintenance and development of the site.

As part of a solution to these problems, PExOD can help to develop the website in a number of ways:

1. PExOD's database is built with PHP, MySQL, and open-source technology. This means that the database can be used to run the website and the pages that the public sees can be customized to suit Dundurn's needs.
2. E-commerce and a shopping cart can be added easily with PExOD (see www.newsociety.com).
3. Information will always be up to date on the website, as it is automatically reflected by any changes made in the administrative side of PExOD.

4. Searchability would be greatly increased, by title, author, subject, or any other identified category thus further enabling e-commerce.

These four features of the program would greatly benefit the existing website and help raise the profile of Dundurn, while also increasing sales revenue.

Opportunity For Website Development

During the final week of my internship an opportunity presented itself that provided the potential to move forward with the integration of PExOD into the website. We submitted a proposal for funding to the Ontario Media Development Corporation for a grant to increase sales and marketing in an area of Dundurn. The plan that was put forth included a complete overhaul of the website, using PExOD as the driving force behind the site, as well as adding e-commerce, new web pages, specialty web pages, and librarian web pages. Our proposal was accepted, and Dundurn received a grant to move forward with the plan in September.

To move forward with this project, Dundurn must transfer the existing PExOD database from the server at SFU to our own domain at dundurn.com. Once this is accomplished, we will have complete autonomy over how the software, additional programming, and template pages are utilized. As PExOD is built on top of open-source technologies, a computer programmer can read and modify PExOD's existing code and/or web-page templates to meet specific needs. This includes building in the capacity for e-commerce and the ability to back up all information.

One goal for the future of Dundurn's use of PExOD and the website is to include all out-of-print books. Dundurn has published about 1200 books, roughly 800 of which

are out of print. However, many customers still inquire after titles that were published twenty-five or thirty years ago. Normally, once these books are no longer in print, they are no longer available to purchase, however with print on demand, this can change. Dundurn has established a relationship with Printorium, a company that prints books (from one copy to as many as you require at a time) either from PDF files, or by scanning a hard copy of a book. We hope to set up the website so that all out-of-print books can be entered into the database and tagged as out of print. Once this has been completed, if a customer selects one of these books on the website, he or she will be informed that it is out of print; but if he or she wishes to find out the costs associated with printing a copy, or several copies, a form can be filled out online. This will allow the staff to monitor which books are asked for frequently, as well as generating once lost revenue as orders for absent books could be filled.

Implementation of PExOD at The Dundurn Group

Within my first week at Dundurn, I manipulated the same bibliographic data that was generated for the fall list at least twelve times. The information had to be formatted in different Excel spreadsheets and sent to Amazon, Chapters, Book Express; transferred to an Access database to be sent to Gazelle; submitted online to Ingram; and written out by hand on forms to the *Canada Books In Print* listing. I was astounded at how many different people and organizations required this information; and that it had to be structured in so many different ways. I grew increasingly frustrated by how long it took to first find out who kept certain pieces of information, then having to track down what I needed, and then spend days filling out different forms and submitting them to the various trading partners. Part of the problem was that Dundurn does not have a central computer server on their network. This means that any bibliographic information is typically sent between employees via email as staff members requested it. This form of exchange generates a lot of waiting time, (plus the time involved at each end) thus decreasing the potential for efficiency in the existing processes at the company.

After several meetings in early May 2004 with Beth Bruder, the Vice-President of Sales and Marketing at The Dundurn Group, we decided that what the company required was a systematic way of organizing data that would meet the required ONIX standards set out by EDItEUR. After writing and discussing a detailed proposal for how this project would operate, the implementation of PExOD at Dundurn commenced on May 17, 2004.

In order to ensure that the project was on track, I generated weekly progress reports that were sent via email to Rowland Lorimer, Beth Bruder, and John Maxwell.

First Steps

The first step in creating a database for Dundurn was to find out where the required information was physically located. To do this I spoke with the head of each department to find out what bibliographic data they used in daily routines, how it was used, and where it was stored. The main source of digital data was from the existing flat file database used for the website. Data was pulled from the site into a spreadsheet where I could manipulate it and clean it up. Though this was the fastest way to gain access to the data for approximately four hundred titles, several different employees had entered the information into this existing database. This meant that some of the data were missing or entered incorrectly and it took about three weeks before I finished entering the missing data, correcting the existing data, and formatting all of the data to be uploaded into PExOD.

Once the data were entered, the second step was to evaluate the fields in PExOD and decide which ones were of use to Dundurn. In order for PExOD to benefit the company in terms of its ability to export information, I knew that the information that was entered initially had to meet to the bronze standard of ONIX information set out by BookNet Canada. Thus, my first priority was to fill out the basic (bronze standard) title information such as ISBN, title, contributor information, price, page count, dimensions, format, biography blurb, and the description for each book. These basic details for a title are just enough information to hold a place in the database for each title as a starting point from which to build a better profile for each title. Once this second decision was

made on which fields to use initially, I could start uploading the data directly into PExOD. This is where I had the most trouble with the implementation process.

Not Just a Click of the Button: Uploading Data into PExOD

As the production schedule at Dundurn was already focusing on the Fall 2004 title list, it was beneficial for the fall title information to be the first that was entered into PExOD. If basic title data (up to the bronze level) are the only data being entered then the ideal way to accomplish this task is to format information about the upcoming titles into a spreadsheet, create an upload profile in PExOD, and then click "upload" so that all of the information flows into the database. While the PExOD interface can also be used, and certainly it would be more efficient to use it for a richer data set, a simple spreadsheet for basic data is the easiest data-entry method. However, uploading from a spreadsheet is not as easy to accomplish as it sounds. At the time, no training manual with the software existed and I had had no formal training on its use. Partly for these reasons I had trouble formatting the data so that PExOD would accept an upload so I entered this information for the fall list manually. With a list of thirty-three new titles, and with several other projects to complete and forms to fill out, this process took just over three weeks to complete. It required considerable support and help from John Maxwell and considerable effort on my part. If all new basic information was entered in this fashion, it would take far too much time, and the database would fail to achieve what it as intended to accomplish: reduce the inefficiencies in the management of data and processes at a publishing firm.

In week four I finally got the upload to work. To upload a file into PExOD, a user can create a spreadsheet in Excel and then save it as a tab-separated text file. Particular columns such as the date and different codes must be set in a particular way in order for PExOD to accept the file as valid data. Once the spreadsheet has been formatted, the user creates a profile in the upload section of PExOD by indicating with numbers which column each ONIX tag correlates with, and checking off a box to indicate that you wish to upload that particular tag. Creating an upload profile takes a couple of minutes, and a user can save the profile to use for future uploads, provided that the same information is located in the same columns of a spreadsheet each time an upload is performed. The information must be formatted precisely as it is set out and coded in PExOD, or the upload will not work.

The upload did not work for me initially because I was trying to upload the information without including all of the code information. For example, to indicate if a contributor has written a book, one must enter the tag "A01," but if they have edited it, one must enter the tag "B01" (these are ONIX standard codes). This information must be included in a column of the spreadsheet for the data to be uploaded into PExOD (other data-sensitive areas of PExOD include the fact that the ISBN cannot have any hyphens, and the data must be formatted YYYY-MM-DD). Once I realized how the database required data to be formatted, all of the information I had been trying to upload or manually enter into PExOD over the first few weeks of this project miraculously flowed into PExOD with once click of a button.

The creation of a training manual (provided to me in June 2004 by John Maxwell), has removed a major barrier to PExOD — the inability to upload the

information easily. The manual removes a stumbling block for the publisher as it helps to determine the exact format for each field. Not knowing this information leads to the loss of great deal of time in setting up the initial uploads. Other proprietary database companies may also have complex uploading processes, but they offer to input all of your data for a company, conduct training sessions, and supply technical support around the clock.

Managing Uploads

Once the upload process started to work, the bulk of the batch imports were imported quite quickly in several small batches. Instead of creating one cumbersome spreadsheet full of data, I divided uploads into segments such as title, publisher, and supplier information. The title information upload includes the ISBN, title, subtitle, contributors, the contributors' roles, page count, book description, dimensions, contributor bio, and publication date. Publisher and supplier uploads contain basically the same information with a few exceptions. The publisher information upload includes the name of the company and contact information while the supplier information includes this information plus the ISBN, price, and currency information for each title. Splitting the information to be uploaded into three separate spreadsheets helped me reduce the possibility of potential human error while also adding a safeguard that all of the information is correct because it is handled in more manageable amounts because there are fewer columns to look at.

After the Upload

Entering data is a time-consuming task. Though the majority of the data PExOD requires can be easily formatted and uploaded, a lot of the title information that can be entered into PExOD was not available at Dundurn in digital form, was missing, or was incomplete. I also encountered other problems. Even though the upload contained author biographies and descriptions of the books, the text was often truncated when uploaded, and the HTML mark-up for the web was stripped out of the text during the upload process. This occurred within the data transfer itself: a result of the database recognizing certain pieces of information and discarding that information which it did not find to be accurate, necessary, or supported within the parameters set up in the upload. For example, it would have helped if I could tell the database that the text was in HTML format. I also did not upload BISAC codes for a title during this process. The BISAC codes are generated for a title as a user clicks on multiple and different variables pertaining to the subject of a title. For instance, a user would choose "Fiction" and then "General" or "Mystery" to further qualify the code. In order for the upload to work, you would have to enter the various BISAC codes, and it is actually much quicker to simply click on the choices than to scroll through a large spreadsheet to find and type in the codes.

I was, however, able to upload the bulk of the information into PExOD so I entered the missing text and HTML mark-up to each title profile directly into PExOD. To do this I copied and pasted the information from Dundurn's existing website and added text to each of the four hundred title profiles. I also entered the BISAC code information, a custom code from a list I created in PExOD[3], and the JPEGs of each cover

image (which have all been resized so that they are all the same dimensions). Many of the covers for the older books were not available as digital files and had to be recreated with the use of a scanner (provided that we had a hard copy in the office).

Another glitch occurred on July 5, 2004 (the eighth week of implementation at Dundurn), when the new PExOD, version 2.20.15 was released. When the upgrade took place, all of the ISBN information for the titles that had already been entered into PExOD disappeared. This was extremely odd because PExOD will not accept a title unless it has this qualifier. Dundurn and Orca were the only two companies that experienced this difficulty. As Dundurn did not start using PExOD right away, they were not included in one of the upgrades that PExOD went through. This meant that the newest upgrade was added to older software than it was created for, thus creating some glitches. Luckily I had only entered about fifty titles at this point, so I quickly re-entered the ISBN information so it would be available during the training sessions that were to take place the same day the ISBNs mysteriously disappeared.

Staff Training and the Publishing Process

The biggest challenge in the implementation process was introducing PExOD to the staff of Dundurn and ensuring that PExOD was implemented and used effectively throughout the publishing process. I developed and delivered five separate training sessions so that each session could be tailored to the different department or person who was being trained. In each training session I introduced the group to PExOD, showed them how to access the administrative interface, and walked them through each of the tabs and fields, stressing those that would be most important to the group being trained.

The staff seemed to be split in the initial enthusiasm or rejection of using a new technology in the company. The editorial staff was the most reluctant, as they did not see how PExOD would benefit them personally. Upon training the editorial staff, I discovered that they would use PExOD mainly for author contact information, and award submission information. The publicity and sales and marketing teams were extremely keen on using PExOD. Both of these positions necessitate the combination of all information about a title as cover images, price, author, and title information are required to create flyers, sell sheets, and press releases for marketing and publicity campaigns.

My main goal throughout the implementation process at The Dundurn Group was to articulate and promote the ways that PExOD would help each individual staff member in his or her daily routine. However, articulating the necessity and benefits of PExOD felt like an uphill battle. Throughout the entire implementation process I constantly communicated and promoted both verbally, and through emails what PExOD could do and how it could change the processes at Dundurn for the better, while also reducing frustration in completing tasks. Every weekly production and marketing meeting brought issues to light where PExOD could alter the existing process in slight, but obvious ways.

From my initial consultations with each department regarding the storage and usage of data, the publishing process at Dundurn became apparent. From the time that a book contract is agreed upon and signed, an ISBN is assigned to the title, the data can continue to be created as the book flows along the production process. With PExOD, all title information can follow and change as the book progresses through the publishing process already established at Dundurn so that the information is always current and

available. The publishing process and how PExOD can influence this process at Dundurn is described in the following table:

Table 1: Impact of PExOD on the Editoral Process at Dundurn

Editorial Process at Dundurn	**Impact of PExOD on the Process**
The Editorial Board meets and decides on the books that will be published.	Contact information for proposed authors can be entered into PExOD.
Acquisition of title and the contract is signed.	ISBN can be assigned and entered into PExOD with tentative title info and author info
Concept meetings to discuss new books with the staff.	Flesh out more of the title information such as the title, subtitle, price, dimensions, page count, and publication date of the new books.
Design of cover (not final) and all title info for catalogue.	Enter information from catalogue for each title (price info, pub date, trim size, page count, illustration info, subject, BISAC code, blurb, and author bio) and upload cover image as a media file (I have been using a 4.5" size) an author image would also be helpful to get at this time (probably 2" in size).
Information sent to on-line booksellers and distributors.	Instead of entering this information and sending it to each distributor separately, exports can easily be created to send the information that has already been entered into PExOD. This export option also allows you to send updated information at a later date.
Receive manuscript from author.	Manuscript can be put into PExOD if needed, or sample chapters can be entered. This would help with completing the tip sheets for the sales conference.
Meeting with reps at sales conference and gaining additional feedback from them.	Changes can be entered into PExOD, cover images are readily available for preparing necessary sales and publicity documents, Contact info is also there.
Substantive editing.	Author contact info will be needed at this stage.
Copy-editing.	Author contact info will be needed at this stage.

Author changes.	Author contact info will be needed at this stage.
Design and layout of book done.	New cover image can be uploaded (if there is one).
Proofread page proofs in house.	Sales and marketing will be using PExOD for title information.
Proofread page proofs by author.	Sales and marketing will be using PExOD for title information.
Changes entered into layout.	Author contact info will be needed at this stage.
Approved by the Senior Editor and Publisher.	Printer contact information (if needed) can be accessed.
Sent to printer.	Finalized page proofs or sample chapters can be uploaded as media files (JPEG or PDF).
Returned from printer.	Final pub date and page count entered into PExOD.
Sales, marketing, and publicity (run parallel to the above throughout this process).	Sales and marketing will be using PExOD for title information and author contact info will be needed at this stage.
Publicity campaigns for a title.	Publicity will need author images, cover images, title info, and author contact info will at this stage.

PExOD can fit into this established process, and can help achieve efficiency within the company by doing so. By using PExOD in this way, it becomes more than just a bibliographic database, but a publishing management tool, a marketing tool and an internal communications tool thus increasing efficiency of the publishing processes at Dundurn.

Impact of PExOD and Logistical Concerns at The Dundurn Group

To determine the impact that PExOD made on the staff and the publishing process at Dundurn, I sent questionnaires (see appendix A) via email two weeks after the training sessions to the ten employees, asking them a series of questions about their experiences with PExOD. Sixty per cent of the questionnaires were completed and returned to me, offering valuable feedback on how the staff views, uses, and would like to use PExOD.

Of the returned questionnaires, sixty-six per cent of the staff had used PExOD since training, and one hundred per cent of them viewed PExOD as a useful and needed tool for the company and useful in all areas. One respondent stated that PExOD "is an essential step in aligning ourselves with key customers. It also increases efficiency as all the appropriate information is in one central location." Another stated

> it's like one-stop shopping. Pretty much everything you need to know about a title you can call up at your desk. Also, it will save me a lot of time if all the cover jpeg's and author photos can simply be uploaded to it so people aren't constantly asking me for images. When I'm in the middle of a season, especially fall, these requests can be really disruptive and take away from my main concern and priority, which is the design and production of upcoming titles.

PExOD became most useful as a central server for all information in the company. It has started to impact the publishing process and other projects greatly at Dundurn as the staff began to accept the new technology and to refer to it for any information first, before having to ask questions as to where an address or cover image might be.

User Concerns For PExOD

Several concerns were also outlined in the questionnaires, training sessions, and throughout the implementation process. Security of the data, who in the company would be responsible for maintaining the data, and the financial support of PExOD were some of the major concerns that employees indicated.

The most serious concern at this point is for the security of the data in PExOD. The way PExOD is currently built, with administrative and public interfaces, means that only account and password holders can access the administrative side of the program; everyone in the company can access the public web pages. I learned early on that the public pages did not have enough information available to the employees at Dundurn. In order for PExOD to be a truly useful tool, it would have to act as a central server for information, and all employees would need a password to gain access to the information entered in the system. This is a problem because, once you have a password for PExOD, you can change, edit, or delete any information that has been entered. Though staff members are not likely to sabotage the company's data-system intentionally, they may delete or enter something in error, or may enter items inconsistently. One way to solve this problem would be to restructure PExOD so that it can handle departmental passwords, or have an option to "view only," so that PExOD can be used to its maximum potential at The Dundurn Group.

As this was a recurring question from the staff at Dundurn, I spent a lot of time thinking about and discussing the possibilities of a password protection plan with John Maxwell. The password protection plan, in theory, would work to alleviate this anxiety, and to better incorporate PExOD into the processes at Dundurn. There are a few different

ways that this principle could be incorporated into PExOD, each of which would create a different set of procedures for how the software is integrated into the publishing process.

A basic password protection plan could be set up in such a way so that each staff member had his or her own user name and password to access the entire administrative side of PExOD. When a user creates a new account in PExOD, there could be a box to click for "view only" rights for the majority of these accounts and the ability to set up administrative accounts for the entire system. This means that one person would take on the role of a central administrator and would control the all of the data in PExOD: uploading it, entering it, and exporting it. The rest of the accounts would be able to access all of the title list information and could view anything within the different profiles, but not the account, upload, or export sections of PExOD.

An alternative way to implement a password protection plan is to incorporate different passwords for each user with a "view only" option for different sections of PExOD. This plan is slightly different from the prior plan as each user would not have access to all of the title details in PExOD, but would only have access to certain sections of the database based on the department in which the user worked. To create this password protection plan, each field and sets of fields in PExOD would need to be assessed and allocated to a certain department. While the different departments are already delineated in PExOD, it would not be adequate to simply group the information in the different tabs without first finding out which department would be responsible for entering the data. Moreover, each publishing house has a different set of processes for managing data, so the stress of customizing PExOD for each individual publisher is

something that simply cannot be a possibility given the relative costs that would be associated with customizing the program for each user.

A third option to work around this problem would be to create a set of web pages, similar to those of the public templates, but which would contain more data from PExOD. For example, while the current public web pages are useful for obtaining cover images, book descriptions, and author bios, they do not include author contact information, print run information, prize, or review information. If another set of templates were created with a slightly different URL, that only the staff can use, then we could pull any and all information from the database to these pages for internal use. This would ensure that the data in the database maintained its integrity, while also providing the staff with the information that they require from the database.

As correlating, distributing, acquiring, and caring for data is a full-time position, it may not be possible for each employee to take the extra time to enter data into PExOD. With different people entering different sections of information, there is more chance of error in the data, which would ultimately slow down the exporting process, as the exporter would still have to find and replace any incorrect data. This plan would require more dedication and communication amongst the staff members, and if one person did not enter the data, the entire project becomes unsuccessful because the company would then degenerate back to its former state of entering data on an as-needed basis. As the staff at Dundurn is already extremely short on time, I believe that a system with a central administrator would work best for their needs.

The fact that PExOD is not privately funded and has been developed at an academic institution led to questions and concerns by the senior staff at Dundurn

surrounding its longevity and quality. With regard to the concern of funding and cost, the PExOD team at the CCSP developed a business plan in August 2004 outlining its proposed manner of operation and the cost to publishers.

Brief Look at the Business Plan

The CCSP's mandate for PExOD is based on "research, teaching, and community service," and the PExOD developers "are convinced that the industry as a whole would be much better off investing in a common solution that readily meets the needs of the majority of BPIDP recipients" (Lorimer et al, 2004: 7). At the same time, the CCSP recognizes that small and mid-sized publishers do not have the financial resources to invest in technology. The proposed manner in which PExOD can be sustained range from private sector control of PExOD to making the code available or inaccessible to selling to a larger entity or industry control with built in leadership (Lorimer et al, 2004: 7). Though all of these are proposed options of dealing with the financial development of PExOD, the CCSP would ideally like to make the licenses for PExOD available to Canadian publishing forms fee of charge, while charging publishers for mandatory installation and support packages (Lorimer et al, 2004: 23). The rationale for charging a mandatory installation and support package is to ensure that users gain basic competency in the software to allow them to take full advantage of its capabilities.

The installation and service packages would depend on the size of the publishing company. Year one would cost $750 for firms with annual revenues of less than $500,000, $1,500 if the revenue was $500,000 to $1 million, and $2,000 for companies with annual revenues greater than $1 million (Lorimer, August 2004: 13). Tech support and maintenance fees would also apply for subsequent years ranging from $150 to $450

(Lorimer et al, August 2004: 13). Additional services and development of PExOD that might come up would be charged on a fee-for-service basis.

A sliding scale of this sort ensures that PExOD can function as a solution for all publishers, relative to their revenue. It also addresses the fear that PExOD cannot develop properly because it does not have adequate funding, and will help to ensure that the project will be able to develop to further meet the varying needs of publishers. This form of funding would help secure the future of PExOD in Canadian publishing.

The Future of PExOD at The Dundurn Group and Conclusion

The creation and storage of bibliographic data are processes that are constantly evolving. The creation of new technologies, standards, and methods aim to improve the process of data storage and exchange within the supply chain. ONIX has been a major step forward for the publishing industry, as it has created the basis from which all other systems can grow. PExOD is one of these systems. PExOD has grown out of the knowledge that small to mid-sized publishers cannot afford to maintain expensive database systems, nor do they have the human resources or time to allocate to data management. This being the case, PExOD provides a solution to publishers who need to keep abreast of the changes in ONIX standards, and the requirements from large retailers, so that they can stay in business. The Dundurn Group is a prime example of this publishing demographic.

In a few short months, Dundurn has transformed from a company with very little organization of bibliographic data to one where all of the information an employee requires for a title is in a central, digital location. Implementing a new technology, such as PExOD was a challenging task, but as the data were uploaded, the staff started to see the necessity of such a technology. PExOD helped to reduce inefficiencies when it came to filling out spreadsheets for trading partners, running the website, and finding information about a book quickly. It also is used as a server for the company, which has become useful in that information is no longer segregated into different departments.

PExOD has become a part of the daily routine for most of the employees at Dundurn. The publicity and sales and marketing departments always have the administrative side of PExOD open, and find that it has helped to increase efficiency and productivity to have all of the data and contact information they require to complete tasks in one, central, digital location. As for its future use in house, PExOD will continue to take on more of a role as more data are entered into the system.

PExOD has also had a positive impact on the flow and exchange of data at Dundurn. For instance, as soon as the information for the spring 2005 list started to become available in August 2004, I was provided with the ISBN, title, and author information. Once this was uploaded into PExOD, I was able to monitor what information was still needed for each title when we started to proceed with the catalogue production. The collection of data at an early stage helped to improve the process for the production of the catalogue, as no one had to scramble to find data in several different places, but could easily refer to PExOD to see all of the available information.

Though some of the staff are forgetful or hesitant in providing data, most have been thrilled to hand any new or updated information to me as soon as they receive it. I no longer have to harass employees for data, but automatically have information such as cover images and book descriptions emailed to me. From these interactions, I suspect that the importance of the database is starting to become clear, and that the majority of the staff recognizes the utility that PExOD has to offer.

PExOD is an important tool for exchanging data in the supply chain. Now that all of the data for Dundurn's titles are entered into the database, an administrator can export initial and updated information to trading partners who accept ONIX feeds. If PExOD

developed to a point where these exports could be scheduled and sent automatically, publishers could rest assured that the correct information is constantly being fed to trading partners, further decreasing the amount of time spent formatting this information.

The future of website development also rests with PExOD. Dundurn plans to launch a new website in January 2005, using PExOD as the database for the site. PExOD allows the searchability of the site to be enhanced easily (with the use of PExOD's title search bar), to allow users to browse by category, and to add a shopping cart to the site with relative ease.

While PExOD serves its purpose as a bibliographic database quite well, a complete publishing management solution for the company would require additional modules. These include managing rights, royalties, accounting, contracts, marketing and publicity campaigns, as well as the ability to pull sales data from CPRESS and create sales reports from this data. As it is built on open-source technology, PExOD can be programmed to fit into existing computer programs that have been built to manage these processes. At first blush, this appears to assume that a publisher possesses or can hire solid IT personnel, and that the company the financial means by which it can support further development of PExOD internally. PExOD's business plan addresses this potential weakness. Groups of companies can form cost-sharing cooperatives to commission additional functionality. Subsequently other community members can benefit directly.

There are several bibliographic data solutions on the market: Klopotek from Germany, ANKO from the United Kingdom, Bookmaster from Australia, Acumen and VISTA from the United States, and Optiwise and HiPoint from Canada to name a few. Of

these programs, PExOD, ANKO, Acumen, and HiPOINT are all stand-alone technology in that the database is hosted on the publishing company's own server. The other programs operate as Application Service Provider (ASP) which means that a user must log in through the web to gain access to the database that is hosted by the software company. All of these products provide the ability to store all bibliographic data, and to export these data in ONIX, and each of them goes beyond this basic function to offer complete publishing management solutions to a publisher (and, in some cases to distributors as well). Klopotek, for example has the ability to generate a media contact list, track the books sent to each contact, and the costs associated with sending out review copies. It generates an invoice, and sends the order for the books directly to the publisher's distributor in a few easy steps. Each of these programs also allows you to generate and analyze daily sales data (brought in from the distributor's data reports), and each can create hundreds of sales reports instantly (currently done manually at Dundurn).

Though these programs can do more with data, one major obstacle to obtaining most of these different programs is that they are extremely expensive, and require yearly administrative, technical support and upgrade fees. That said, it appears that if Dundurn found another cost-effective program that could perform tasks above and beyond that which PExOD currently performs, we would likely move the data to another system; one that further reduces inefficiencies and redundancies in the company beyond the recording and exchange of bibliographic data. As the seed of digitization has now been planted, I believe that Dundurn will constantly look at improving its processes. Whether PExOD can put in place the means for it to work with Canadian publishers to develop additional functionality as firms are ready to adopt it remains to be seen.

The book publishing industry is rapidly moving towards systems of digitization and automation when it comes to bibliographic data management. PExOD is one of the many tools available to help publishers cope with these changes in an affordable manner. ONIX is a relatively new development and the processes by which data needs to be managed and the standards by which it is required to flow throughout the supply chain are not completely certain yet. However, as these standards and processes continue to evolve throughout the industry PExOD is a valuable tool for publishers to use, as the program is always updated to remain current with the standards set out by EDItEUR and BookNet Canada. I believe that PExOD will continue to evolve as conditions change, and that publishers will be able to rely on this technology as an accountable database system built by those who truly have the needs of Canadian publishers at heart.

Appendix

PExOD Questionnaire

I am sending you this questionnaire as a follow-up to our training session with PExOD. Please fill out the form and email it back to me, or if you are more comfortable printing it out and providing me with a hand-written form, that would be great too. Could you please send this questionnaire back to me at hsanderson@dundurn.com by August 6 at 1pm. Thank you for your time and assistance with this project.

1. Over the past few weeks have you used PExOD? Yes ____ No ____

2. If "Yes," what tasks have you used PExOD to perform? If "No," why have you not used PExOD?

3. What do you use PExOD most frequently for?

4. What advantages/disadvantages have you encountered by using PExOD?

5. What role do you see PExOD filling in your department? (Please state department).

6. What role do you see PExOD filling within The Dundurn Group?

7. Are there certain tasks that you want PExOD to perform which it does not currently perform? Please explain.

8. Do you find PExOD to be user-friendly? Please explain.

9. Do you see PExOD as a useful tool for The Dundurn Group? Please explain why/why not.

10. Do you have any concerns about using PExOD as the database for The Dundurn Group? Please explain.

Please state any additional comments

Reference List

Barnes, Roger and Rowland Lorimer. "Book Purchasing in Canada: 1997 In-Store Survey," Canadian Publishers' Council: Toronto, March 27, 1998.

Book Industry Study Group. "ONIX." Book Industry Study Group Home Page. 31 August 2004. <http://www.bisg.org/onix/index.html>.

BookNet Canada. Bibliographic Standard Guidelines and Appendix, June 22, 2004, Version 1.3.

BookNet Canada. "Bibliographic Data: ONIX." BookNet Canada Home Page. 25 August 2004. <http://booknet.torque.net/booknet/biblio_onix.shtml>.

EDItEUR. "ONIX for Books." EDItEUR Home Page. 25 August 2004 <http://www.editeur.org>.

Evans and Company. "A Review of the Canadian Book Industry," Toronto: Association of Canadian Publishers, 2001.

Lessig, Lawrence. Code: And Other Laws of Cyberspace, New York: Basic Books, 1999.

Lorimer, Rowland and Associates. Build It and They Will Flow: The Distribution System for English-Language Books In Canada, Vancouver, BC, September 2003.

Lorimer, Rowland, John Maxwell, Doug Plant, and Craig Riggs. PExOD: The Publisher's Extensible ONIX Database Business Plan, Canadian Centre for Studies in Publishing, Vancouver, BC, August 13, 2004.

Lorimer, Rowland. Vibrant But Threatened, Canadian Centre for Studies in Publishing, Vancouver, BC.

MacSkimming, Roy. The Perilous Trade: Publishing Canada's Writers, Toronto: McClelland & Stewart Ltd., 2003.

Neilson Book Services Ltd. "About Book Data," BookData Home Page. 2004. 5 October 2004.<http://www.bookdata.co.uk/about_book_data.html>.

Sasmor, Evelyn. Computers in Libraries, February 2004, Vol. 24 Issue 2, 25.

Statistics Canada. "Profile of book publishing and exclusive agency in, for English language firms." Statistics Canada Home Page. 2 October 2004. 4 October 2004 <http://www.statcan.ca/english/Pgdb/arts02.htm>.

St. Laurent, Simon. XML: A Primer, Second Edition, California: MIS Press, 1999.

Endnotes

---

[1] Statistics Canada totals for French and English titles in 2000-2001: 113,400 in print overall, 15,707 printed, and 12,002 reprinted

[2] BPIDP was brought forward in 1986 by the Department of Communications Minister, Marcel Masse (Lorimer, Vibrant: 84).

[3] Within the administrative interface of PExOD there is a place to create custom categories for titles which allow a user on the public interface to browse all of the titles by category.