

# **The Economics of Digital Access: The Early Canadiana Online Project**

Bruce R. Kingma  
Associate Professor  
School of Information Science and Policy  
Department of Economics  
Interdisciplinary Ph.D. Program in Information Science  
University at Albany, SUNY

135 Western Ave  
Albany, New York 12222

[bk797@albany.edu](mailto:bk797@albany.edu)

[www.albany.edu/~bk797](http://www.albany.edu/~bk797)

---

© 1999 DRAFT REPORT 08/10/99 2:00 PM. Do not copy, print, cut, or reproduce without the written permission of the author. All rights reserved. Use of any material, in whole or in part, is expressly forbidden without prior written consent. I would like to thank Pam Bjornson, Meredith Butler, Marshall Clinton, Malcolm Getz, Tim Neff, Guy Teasdale, and Karen Turko for their assistance with this report.

## Contents

Executive Summary	3
Introduction	4
An Economic Theory of Digital Access to Information	5
Patrons of Early Canadiana Online	5
Producers	8
Intermediaries	9
Individual and Shared Costs and the Role of Information Intermediaries	9
Market Forces: Demand and Supply	10
Cost Estimates of Early Canadiana Online	10
The Cost of Print	11
The Cost of Microfiche	12
The Cost of Digital	15
Economies of Scale	17
Cost of Access to Digital Information	19
User Costs of Access	20
Pricing of Digital Information in a Networked Environment	22
Conclusions	25
Appendix	26

## Executive Summary

This project examined the economics of the production, storage, and distribution of information in print, microfiche, and digital format. The costs of these formats are estimated as part of the Early Canadiana Online project which digitized over 3,000 titles of the Canadian Institute of Historical Microreproductions collection. The results include:

- The average cost of producing a 216 page book in digital format from microfiche is \$258.82 (C\$) plus any copyright fees. The annual cost of storage and access is \$39.20 per book. In theory, these costs can be shared by the number of libraries and patrons that access the digital copy over the Internet significantly lowering the costs per library.
- The average cost of producing a book on microfiche is \$333.11. Given the number of number of copies sold by CIHM, the cost per library is \$9.69-\$13.56 per book. The annual cost of storage and access in a university library microtext room is \$0.16 per book.
- The cost of a book in print format is the purchase price of a book. The annual cost of storage and access is \$5.89 in a rare book library.

These cost estimates show that digitization of texts can provide significant savings if shared by a sufficient number of subscribing organizations. Networked access to digitized texts also provides several economic benefits to users including: 1) increasing the availability of these texts to patrons of organizations with access to the Internet, 2) decreasing the opportunity cost of patrons' time spent accessing digital copies rather than traveling to libraries to use print or microfiche copies, and 3) providing electronically searchable texts making it easier for users to find items of interest. These increased benefits should result in a significant increase in use of the digital information relative to use of the print or microfiche copies.

Previous studies have estimated the marginal costs of production, acquisition, and storage of books, microfiche, and digital copies of texts. This study includes all costs associated with the production, cataloging, and sales of texts in microfiche or digital format. Therefore, the estimates of the cost per book in print, microfiche, or digital format are average cost estimates based on digitizing over 3,300 titles and 650,000 images. Individual libraries engaging in small digitization or microfiche projects may have lower costs per text but the final product may not be of a quality needed for national or international sales. Large-scale projects that include cataloging and sales of several thousand texts are likely to experience a similar cost structure as estimated in this study.

Questions remain about financing the digitization of texts. Revenues can be collected from subscriptions, donations, sponsorships, and grants to produce, archive, and provide access to the digital texts. Economic theory suggests, however, that pay-per-view or pricing access for individual subscribers should be avoided.

## Introduction

Digital texts in a networked environment hold the promise of lower cost access to information by a greater number of users than can be provided by printed texts. Projects such as The Making of America<sup>1</sup>, Project MUSE<sup>2</sup>, JSTOR<sup>3</sup>, and the Early Canadiana Online<sup>4</sup> project investigated in this study offer access to digital texts over the Internet to millions of potential users. These digital projects also offer the promise of lower costs by avoiding the cost of printing and shipping multiple copies of a text for patrons. In theory, once the fixed costs of digitization are incurred there is a zero marginal cost of providing an additional copy.

The potential benefits of digital access are considerable. Patrons, who previously traveled to a repository of rare books or a microfiche room at a research library, can instead access historical information from their desktop. This dramatically decreases the time and effort patrons spend traveling to the source of the information. This also dramatically increases the potential benefits to new patrons who can now access historical texts that previously were only available at sites too distant for them to consider. The economic question is whether the cost of digitization is lower than this future stream of benefits.

This study examines the economics of digital, microfiche, and print access for the Early Canadiana Online project. These three methods include the costs of archiving and providing access to original print materials, microfiche copies of these materials, and digital copy accessible over the Internet. This study examines the production and storage costs, opportunity costs to patrons, and potential market for digital access to the Early Canadiana online collection.

Data collected for this study will be important in determining the level of investment for future digitization projects of historical materials. Previous studies by Anne Kenney<sup>5</sup> at Cornell University and Paul Conway<sup>6</sup> at Yale University investigated the costs of digitization. However, both studies measured the marginal costs per image of primarily in-house scanning. This study includes all costs associated with the production, cataloging, and sales of texts in microfiche or digital format. The cost estimates in this study are considerably higher than the marginal cost estimates in previous studies but are a more accurate estimate of the full costs of the production of microfiche or digital projects from start to finish.

This study also investigates the benefits of digitization. The primary benefit of these digital projects is the return to patrons from accessing these materials. Once digitized,

---

<sup>1</sup> <http://www.umdl.umich.edu/moa/>

<sup>2</sup> <http://muse.jhu.edu/>

<sup>3</sup> <http://www.jstor.org/>

<sup>4</sup> <http://www.canadiana.org/>

<sup>5</sup> Anne R. Kenney, "Digital to Microfilm Conversion: A Demonstration Project 1994-1996"

<http://www.library.cornell.edu/preservation/pub.htm>

<sup>6</sup> Paul Conway, "Yale University Library's Project Open Book" D-Lib Magazine, <http://www.dilib.org/dlib/february96/yale/02conway.html> February 1999.

stored, and made accessible over a campus network or the Internet the materials are more easily accessible to more patrons. Patrons, who previously had to travel to a library with the original or microfiche copies of the materials, can now view them online from their home or office. Analysis of the data collected on use of the digital images, microfiche, and original texts will be helpful in predicting the use and benefits of other digital projects of historical materials. This will enable researchers to determine the return to investment of future digital projects.

### **An Economic Theory of Digital Access to Information**

Digitization of information provides lower costs than print products for the production, distribution, and access to information for producers, consumers, and intermediaries. Digital access results in on-demand access to information for patrons or consumers, lowering the opportunity cost of access. Consumers can more easily view digital information over networks without having to spend time traveling to the library. The low cost of web development and word processing lowers the cost of producing information in digital form. Producers do not have to print and distribute copies but can instead mount digital products on a local server enabling network distribution. Likewise, digital access provides lower cost distribution by intermediaries such as libraries, saving the costs of storing and circulating printed materials.

The Early Canadiana Online project includes all three of the stakeholders in the production and consumption of information. Libraries with rare book collections such as the University of Toronto Library and the Laval University Library provide access to original print texts. These libraries also provide access to the Canadian Institute for Historical Microreproduction's microfiche copies of these print materials. In this instance, CIHM is the producer of the information while the libraries are intermediaries in providing access to patrons. Patrons of this information include students and faculty accessing the CIHM collection whether in print, microfiche, or digital form.

#### *Patrons of Early Canadiana Online*

With the creation of the Early Canadiana Online patrons have three possible methods for accessing this information: digital, fiche, or original copy. Patrons incur a cost of access depending on their choice of method of access. These costs can be divided into fixed and variable costs. Variable costs are costs incurred each time information is reproduced or retrieved. Fixed costs are costs incurred regardless of the number of items retrieved.

A patron's choice of access will depend on which method provides lower total costs. A patron viewing images from a single text may have lower fixed and marginal costs in using the print than in using the fiche or digital formats. Accessing the print may require only travel to the library, selection of the text, and turning the pages. There are no learning costs or costs of expensive machines or network connections when using print. Accessing the fiche may require travel, selection, and determining how to use the fiche. Accessing the digital format requires the use or purchase of a computer with a network connection.

Microfiche or digital access is more likely to have a lower total cost when more than a single text is used. While multiple texts may be found in the same library, the fiche collection may contain items not found in the library's print collection. Accessing these items from another library would require the patron to incur an additional cost of traveling to a second library. Learning how to use the microfiche collection is likely to have a lower total cost than traveling to more than one library in order to use the needed items in print form. Likewise, digital access may provide access to more images at a lower total cost than fiche or print.<sup>7</sup>

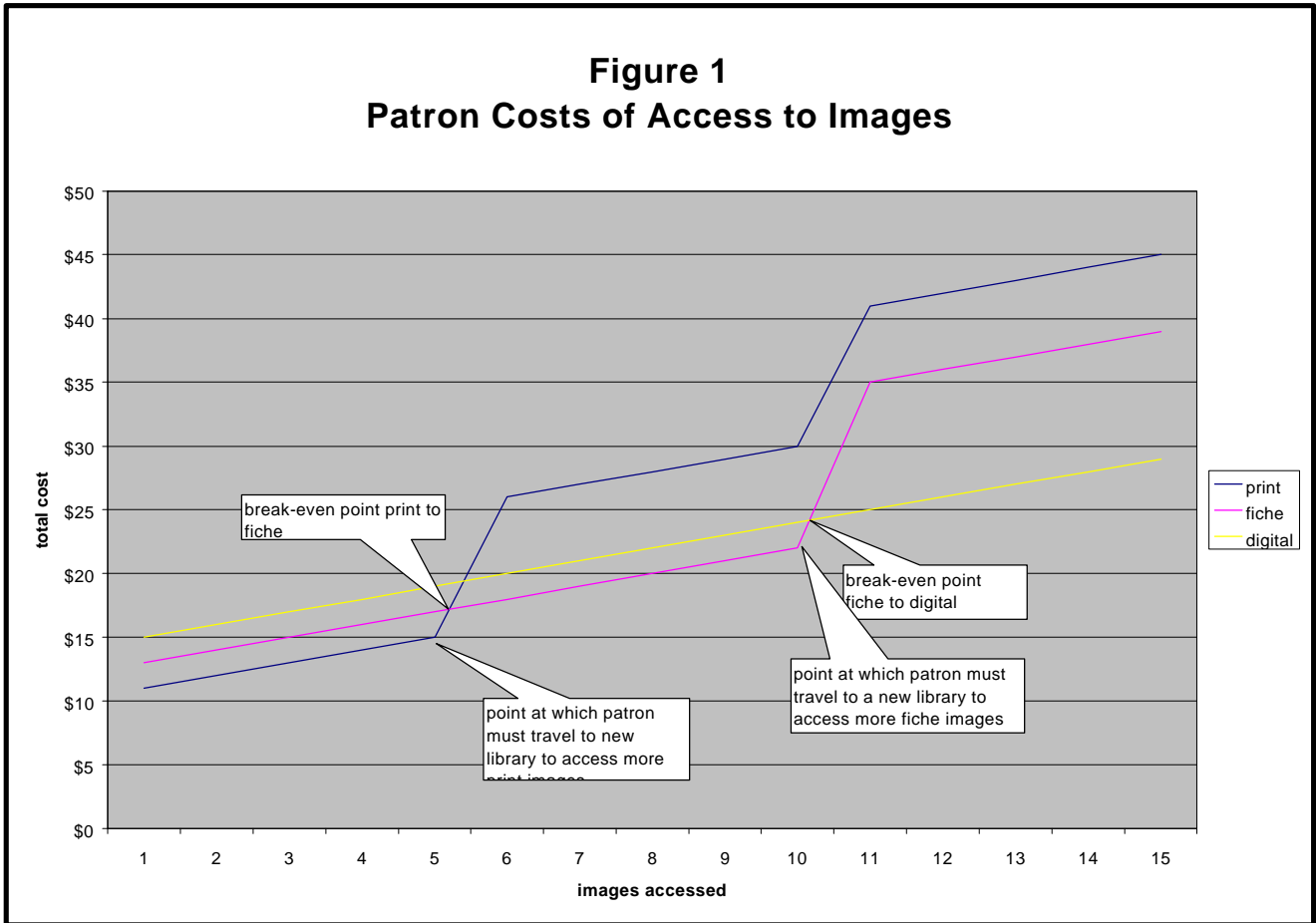
<i>Patron Costs of Access</i>		
<i>Method of access</i>	<i>Variable costs per image retrieved</i>	<i>Fixed costs per set of images retrieved</i>
Print	Minimal, possible costs of photocopying, possible time costs of retrieval for each new item.	Travel to archive for original texts, possible wait for text to be delivered by archive staff. Fixed costs must be repeated when different text is needed from a different library.
Microfiche	Minimal, possible photocopy costs, lower time costs for retrieval of new items than with print.	Travel to library, learning costs of using fiche instead of original copy, retrieval time of fiche. Some costs must be repeated if library does not have full set of fiche.
Digital	Minimal, lower costs of retrieval than with print or fiche, lower costs of photocopy since image can be saved in digital form.	Cost of travel to public access terminal or cost of personal computer, browser, and network connection; learning costs of using digital instead of original copy.

Figure 1 illustrates total patron costs for access to information in the three formats. Figure 1 assumes that the fixed cost of digital access is greater than the fixed cost of fiche which is greater than the fixed cost of print. Figure 1 also assumes the number of images available in digital form is greater than the number available in fiche which is greater than the number of print images.

The break-even points in figure 1 represent the levels of use such that two methods of access have the same total cost. Initially a patron will have a lower total cost from print texts. As use of images increases and the library's print collection is exhausted, a patron

<sup>7</sup> At this time the Early Canadiana Online Project provides access to a smaller set of images than the fiche collection. Since only a limited set of the fiche is available in digital form, patrons with access to the complete fiche collection will have greater access to images than if they were to access the digital images using a web browser. However, patrons without access to the fiche or with access to only a limited subset of the complete CIHM collection will have greater access to images in digital format.

must incur the additional fixed costs of traveling to another library. As use increases, the total cost of fiche access is lower than the total cost of print access. As use continues to increase, the total cost of digital access becomes lower than the total cost of fiche access.



If figure 1 accurately reflects the fixed and variable costs of access then digital access is more likely by high end users that require more access to more images. These users incur a high total cost of access to the digital copy but gain greater access to more information. Patrons desiring only a few images from a single text are more likely to look at the original, print copy if it is available in their library. While mid-level users are more likely to use the fiche.

However, what may be an inaccurate assumption in figure 1 is that digital access has higher fixed costs and equal marginal costs to fiche or print. The fixed costs of digital access include the learning costs patrons unfamiliar with digital copy must spend, as well as the costs of having access including a personal computer with network access. Once these costs are incurred patrons may have a fixed cost of digital access less than the fixed costs of fiche or print access. Patrons can also avoid the fixed costs of traveling to the library if they have at-home or office access to the network. The marginal costs of digital access may also be less than print or fiche. Patrons familiar with digital access are less

likely to print materials, instead saving electronic copy on a disk or drive. If the marginal cost and fixed costs for digital access are lower than for print and fiche then digital access will have a lower cost at all levels of access. Patrons will only access the information online.

### *Producers*

The initial promise of digital information was that production costs would decrease when the costs of printing and distribution were replaced in the networked environment. These lower costs have led to an increasing number of free electronic journals published by faculty at colleges and universities. However, there are also costs in digital production. HTML programming costs, patron service costs, and production in both print and digital formats can increase the costs of production. Traditional print publishers have found that additional costs of production are necessary to publish a journal in both print and digital format, increasing subscription costs to libraries that require access to both forms.

Digital costs are lumpy, with a large fixed cost of production, and zero marginal cost to produce an additional digital copy over the Internet. However, digital copies have the same, if not greater, patron service costs as print and microfiche. Patrons need service in any environment. In the digital environment patron services include the cost of server maintenance, the cost of updating the web pages, and the cost of answering electronic mail from patrons that are having difficulties with access. Since the networked environment allow more patrons to access the information than at the library, the cost of patron service may be greater for the digital information producer. Unlike print publications that are produced and then sent to information intermediaries, customer service in the digital environment requires the information producer to provide direct service to patrons.

Pricing in the networked environment can also be a difficult problem for information producers. Classic economic theory would indicate that the price of access should be set equal to the marginal cost of zero to achieve economic efficiency. However, a zero price does not allow for information producers to recover the costs of production. Access to digital products will be sold above the marginal cost of reproduction in the same way that books, journals, and other print products are sold above the marginal cost of an additional copy. This pricing is based on the value of the information good to consumers, rather than the cost of providing an additional copy. Value based pricing in the networked environment is necessary to recover the costs of production.

The role of the library as intermediary is critical in the pricing of information. Libraries purchase information materials and provide access to patrons typically without an access fee. Patrons efficiently use the information since there are no marginal costs in the networked digital environment for providing the information to them and they are not charged for access. The charge to libraries covers the cost of production of the information while the absence of a charge for patrons insures economic efficiency.

### *Intermediaries*

Libraries serve a crucial economic role as intermediaries in the distribution of and access to information. Libraries serve as a point of collective demand for information products providing access to information as a public good to patrons.

The economics role of the library as an information intermediary is to estimate the collective demand of patrons and purchase and provide access to information goods. The collective value of any information product in a library is the sum of the value or benefit all patrons receive from it. This can be estimated by the number of times the information product is used multiplied by an estimate of the benefit from each use. If this collective value exceeds the purchase price then it is economically efficient for the library to purchase it and provide access to patrons. Providing access to an additional patron typically has no cost and therefore should be provided for free to insure economic efficiency.

For digital products there are two possible benefits to library patrons. If the library does not subscribe to the print or fiche copy of the information, then patrons benefit by accessing information previously not available. If patrons have access to the fiche or print original and access to the digital copy is available over the Internet or campus network then the benefit of access to the digital copy is equal to the value of time saved from using the digital copy from the home or office instead of the fiche or original at the library.

### *Individual and Shared Costs and the Role of Information Intermediaries*

The costs for information products and services can be categorized into private and shared costs or the costs of individual demand and public demand for a good. Private costs are the costs to an individual or consumer of his purchase of a good or service. Private costs include the costs of a personal subscription, personal home computer, photocopying papers, and downloading and printing information from the Internet. Shared costs are the costs of information products purchased for public use. Shared costs include the costs of library goods and services. The costs of library goods and services are shared among patrons through tuition payments, tax revenue, membership fees or other sources of revenue.

Information intermediaries also have what can be considered private and shared costs. A subscription to a print or electronic database can be considered a private cost to the library although it is a shared cost to the library's patrons. The fixed costs of producing the database or print journal that several libraries subscribe to is a shared cost among the subscribing libraries. Each subscribing library pays for a share of the fixed costs of production.

The costs of digital information in a networked environment are shared. On the Internet, the costs of reproduction and distribution are zero. All costs associated with digital information in this environment are the fixed costs of producing and storing the

information. These costs are, by definition, shared among the patrons or information intermediaries that purchase access.

### *Market Forces: Demand and Supply*

Digital information in a networked environment results in lower costs of reproduction and distribution for producers and intermediaries and lower opportunity costs for users. However, while digital economics offers lower cost access to information these lower costs also contribute to an increase in the supply of information. Lower costs result in more producers providing more methods of access to more information.

Producers that provide access in print and digital form have higher costs than if they produced print products only. Digital costs duplicate rather than replace print costs. It is rare that publishers of print products cease print production for digital.

Lower costs also mean new information products are produced. New publishers including universities, libraries, faculty and students find that they can produce web-based journals using low cost desktop publishing tools. This has led to an explosion in the supply of new electronic journals.

Libraries are confronted with this ever increasing supply of digital information products. Although the cost of many new electronic journals is relatively lower than print journals, the total costs of subscribing as well as the staff costs involved in cataloging these journals can dramatically increase a library's total costs. Patrons find that the opportunity cost of accessing any given source of information has declined, while the overwhelming increase in the number of information products results in more time being spent on digital information than was spent consuming print products. The digitization of an information product previously made available only in print or microfiche can lower the cost per unit of production, the cost per unit for subscription by libraries, and the cost of access to the information by patrons, while at the same time dramatically increasing the supply of information products. This increase in the number of information products results in substantially higher total costs of production, subscription and access to information.

### **Cost Estimates of Early Canadiana Online**

Estimating the costs of digital projects is necessary to determine efficient investments in digitization of print or microfiched information products. The primary goals of this project are to estimate and compare the costs of three methods of information delivery; print, microfiche, and digital. Data from the University of Toronto, Laval University, and the Canadian Institute for Historical Microreproductions was collected to estimate these costs. Data on the cost of construction of a new electronic library at the University at Albany was also collected.

One significant contribution of the cost estimates in this paper is that average costs are estimated for the production, storage, and use of information in print, microfiche, and

digital formats. Previous estimates have either focused on one type of cost—production, storage, or use—or one format—print, fiche, or digital. In this paper all costs for each format are included. This gives a more accurate estimate of the full cost of information.

### *The Cost of Print*

Table 1 shows the cost estimates for book storage and access. These costs are based on the cost of the Thomas Fisher Rare Book Library at the University of Toronto. Construction costs are based on the 1999 library construction project at the University at Albany. Special environmental controls used in a rare book library imply that the construction costs in table 1 may underestimate the actual construction costs. All costs are shown in Canadian dollars.

	Cost	Cost per volume (cost/volume)	Cost per use (cost/use)
Construction, utilities & maintenance	\$1,586,056	\$3.17	\$ 72.51
Salaries	\$1,105,031	\$2.21	\$ 50.52
Equipment & supplies	\$ 255,799	\$0.51	\$ 11.69
TOTAL	\$2,946,885	\$5.89	\$134.72

The cost of construction, utilities, and maintenance is comparable to an estimate of \$4.68CD from Bowen.<sup>9</sup> However, the cost per use of \$134.72 is significantly higher than the \$1.50CD cost of retrieval reported by Bowen, the \$3CD for the New York Public Library and \$6CD for the Harvard Depository Library reported by Lesk, or the \$9CD maximum retrieval cost estimated by Getz.<sup>10</sup> In table 1 the cost per use is derived by dividing the cost by the number of requests for the year. This inflates the cost per retrieval by adding the costs of storage into the equation. However, it is important to note that the “service” of a library is the use of its materials. All costs when divided by the use of those materials gives an average cost for service which will be higher than separating out only part of these costs for retrieval.

For a comparison with the estimates cited by Bowen, Lesk, and Getz, an estimated 80% of salaries at the Thomas Fisher Rare Book Library is for access. Taking 80% of salary

<sup>8</sup> Costs of utilities, maintenance, salaries, equipment and supplies are based on University of Toronto cost estimates. Construction costs are based on new library construction at the University at Albany. All costs are in Canadian dollars. Exchange rate used is 1.5257. Costs are amortized using 5% rate of interest and life spans as follows: construction 25 years, equipment 5 years, and computers 3 years. Cost per volume based on library capacity of 500,000 volumes. Cost per transaction based on 21,874 transactions.

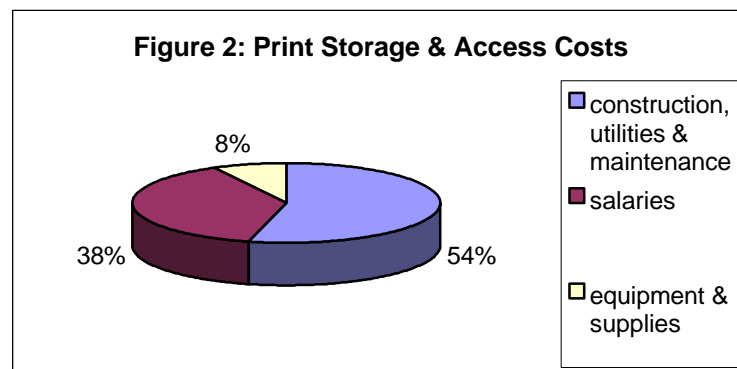
<sup>9</sup> William Bowen, “JSTOR and the Economics of Scholarly Communication” *The Economics of Information in the Networked Environment*, Meredith Butler and Bruce Kingma, eds., Haworth Press 1998.

<sup>10</sup> Michael Lesk, “Substituting Images for Books: The Economics for Libraries” <http://community.bellcore.com/lesk/unlv/unlv.html> April 4, 1998. Malcolm Getz, “Evaluating Digital Strategies for Storing and Retrieving Scholarly Information,” *Economics of Digital Information: Collection, Storage and Delivery*, Sul H. Lee, ed., Haworth Press 1997.

costs yields an estimate of \$40 per transaction for labor, still significantly higher than other estimates. However, a rare book library has concerns of preservation that require additional staff care and monitoring for patron access. In addition, this estimate includes the total cost of administration, vacations, and benefits for employees rather than the marginal cost of retrieval based on a staff member's time spent multiplied by his salary.

Table 1 does not include the cost of purchasing a book. This is important although it will be a small percentage of total costs once the purchase price is amortized over the expected life of storage and use of the book. For example a rare book that costs \$500 but is expected to last 100 years in storage has an annual cost, when amortized, of \$4.80. Table 1 also does not include the value of the land. This can be significant but is different depending on the location of the library.

Figure 2 shows the categories of rare book annual storage and access costs as a percentage of total costs. As expected, the largest component of total costs for books is the cost of space to store them.



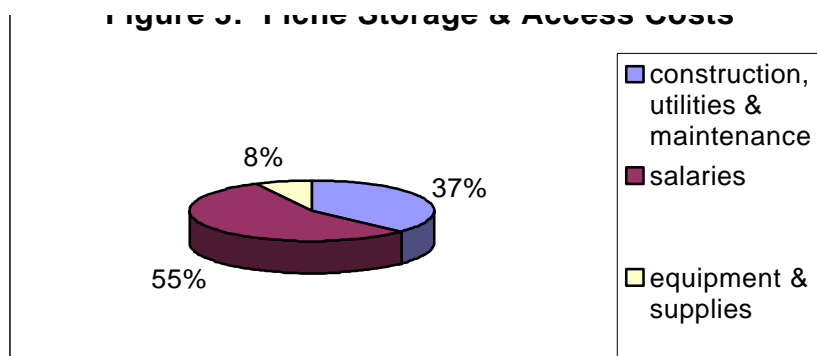
### *The Cost of Microfiche*

The annual costs of microfiche storage and access at the University of Toronto are shown in table 2. Cost per volume is based on a 216 page text, the average size of a text digitized in the Yale Open Book Project.<sup>11</sup> As with table 1 these costs represent the average cost per unit for storage or access. Just as the cost of purchasing a book is not included in table 1, the cost of purchasing the microfiche is not included in table 2.

Both the cost of storage per volume and the cost per use are significantly lower for microfiche than for rare books. This is not surprising since microfiche is intended to provide access to and storage of information at a lower cost than print.

The cost per use is derived by dividing the total costs of microfiche storage and access by total use. As with table 1, this assumes that the value of microfiche storage is for access to patrons. If salaries and equipment are the only costs for access, and 80% of salaries are

<sup>11</sup> <http://www.dilib.org/dlib/february96/yale/02conway.html>



for access, then the cost per transaction can be estimated as \$3.75, which is comparable to estimates of the costs of book retrieval. Both retrieval functions are similar in that staff must locate, check out, and reshelve the requested materials.

	Cost	Cost per volume (cost/volume)	Cost per use (cost/use)
Construction, utilities & maintenance	\$ 170,527	\$0.06	\$2.71
Salaries	\$ 251,602	\$0.09	\$4.00
Equipment & supplies	\$ 34,423	\$0.01	\$0.55
<b>TOTAL</b>	<b>\$ 456,552</b>	<b>\$0.16</b>	<b>\$7.26</b>

Figure 3 illustrates that microfiche costs are more salary intensive. Salaries constitute a larger percentage of the costs of microfiche than in the case of rare books. Rare books take up more space and therefore have a higher percentage of costs in construction, utilities and maintenance.

Table 2 does not include the subscription price of the microfiche to the library. These costs are part of the economic cost of producing microfiche and are shown in table 3. To avoid double-counting, these costs are not shown in table 2. The costs of microfiche production are shared costs. Library subscription fees, grants and donations are used to jointly finance the production of the microfiche as a public good.

Table 3 includes all economic costs of microfiche production including the value of space CIHM uses at the National Library of Canada. While this space is donated to CIHM, it still represents an economic cost of producing microfiche. As with previous tables, the average cost of production is derived by dividing total costs by number of units.

	Cost	Cost per fiche	Cost per image	Cost per volume
Master Copies	\$ 150,000	\$ 16.07	\$0.22	\$ 46.88
Salaries	\$ 602,932	\$ 64.58	\$0.87	\$188.43
Equipment & supplies	\$ 125,880	\$ 13.48	\$0.18	\$ 39.34

<sup>12</sup> All costs are amortized as in table 1 with the exception that amortization for microfiche readers used a 15 year life span. Volumes are considered to contain 216 images. Images are page images. Each microfiche image has two page images. University of Toronto microtext use was 62,856 in 1997 for 3,387,777 units stored in a room of 810 square meters.

Construction, utilities & maintenance	\$ 187,066	\$ 20.04	\$0.27	\$ 58.46
TOTAL (shared costs)	\$1,065,878	\$114.17	\$1.54	\$333.11
cost of microfiche reproduction & sales	\$ 236,092			
TOTAL COST	\$1,301,970			
Total cost per library (30-42 copies)	\$43,399-\$30,999	\$4.65-\$3.32	\$0.06-\$0.04	\$13.56-\$9.69
Annual cost per library (30-42 copies)	\$2,254-\$1,487	\$0.22-\$0.16	\$0.01-\$0.00	\$0.65-\$0.46

The first four rows of table 3 show the cost of producing master copies of microfiche. The cost of producing master copies of microfiche is \$114 per fiche, \$1.54 per image, or \$333.11 per 216 page volume. This is the cost of producing a set of master copies that are then used to produce additional microfiche copies for distribution to subscribing libraries. The cost of the master copies is a shared cost for all subscribing libraries.

If we compare the cost per volume of creating and storing a master microfiche copy relative to creating and storing a print copy, microfiche is expensive to create but has significant savings in storage (\$0.16 per volume per year) relative to print (\$5.89). However, at an annual savings of \$5.73 per year, it would take over 50 years to cover the cost of creation (\$333.11) if the master copies were created solely for the use of one library.

Microfiche is produced by CIHM, not to have a single copy, but to provide multiple copies to libraries that would not otherwise have access to early Canadian literature. With a limited number of print copies, microfiche becomes a cost-effective alternative for providing access. CIHM produces several copies of each microfiche to sell as subscriptions for libraries throughout Canada, the United States, and the rest of the world. By purchasing a subscription, these libraries share the costs of the original microfiche production.

CIHM produces about 30 copies each year for library subscriptions and additional copies of individual microfiche at an additional cost of \$236,092. The last two rows in table 3 show how these costs can be shared among the subscribing libraries. If the full cost of microfiche production is averaged over the 30 copies, the cost of annual production is \$43,399 per library. This includes the shared costs of production plus the costs of making copies. If an additional 12 copies of each fiche, on average, are sold the average cost is \$30,999 per library.

The average cost per fiche, per image, and per volume for 30-42 copies are shown in the final three columns of table 3. The sharing of the full costs of production among subscribing libraries reduces the cost to \$0.04-\$0.06 per image or \$9.69-\$13.56 per volume. This compares favorably to the cost of each library acquiring a printed manuscript. At an annual savings of \$5.73 per volume for each library, it takes 1.7-2.4 years for the microfiche to cover the costs of creation (\$9.69-\$13.56).

Once produced, it is anticipated that a microfiche copy of a text will last for 100 years. The purchase of microfiche is an investment in an archival copy of materials that is expected to provide access to patrons to the information for many years. If the cost of the microfiche is spread out or amortized over a 100-year period, then the annual cost of microfiche production is only \$0.65-\$0.46 per 216 page volume per year. When this is added to the cost of storage from table 2, the annual cost comes to \$0.81-\$0.62 per volume per year for producing, storing, and providing access to a text in microfiche format.

These costs indicate that when microfiche is produced in large numbers to accommodate several libraries, it costs significantly less to produce, store, and provide access to microfiche than to books. This shared cost per library declines further if the number of libraries acquiring subscriptions increases. In addition, the CIHM microfiche subscription provides access to a larger collection of texts than is likely to exist in any single library of rare books. These cost estimates show that microfiche is the more cost-effective alternative to library storage of print to provide patron access to out-of-print texts.

### *The Cost of Digital*

The previous section showed that microfiche is a cost-effective alternative to print. Digitization of texts may be able to provide even greater savings relative to microfiche and print. Unlike print and microfiche, which must be produced and delivered to a library, digital texts have the advantage of being stored remotely but accessed globally via the Internet. The cost of reproduction and distribution of digital information in a networked environment is zero. The only costs are the one-time fixed costs of producing and the annual fixed costs of storing the data. These fixed costs can be shared by the subscribing libraries. In theory this could drive the cost per library to a significantly lower level than with microfiche.

In the Early Canadiana Online Project microfiche was converted to digital format. Microfiche was sent to Preservation Resources for scanning and the University of Michigan for optical character recognition.<sup>13</sup> Cost estimates shown in table 4 are based on contractual costs for scanning and OCR.

The total costs for production are \$236.08 per title or \$1.20 per image. Costs in the second and future years for digital storage and access are \$35.76 per title or \$0.18 per image. This includes the cost of salaries for maintaining the ECO Project database and annual costs of hardware storage. Although the cost of producing digital copy from fiche is less than the cost of microfiche, the cost of storage and access for digital, in this project, is more expensive. This is the result of costs averaging over a smaller number of available digital images which will be higher than the average cost per fiche in a university microtext room which contains hundreds of thousands of microfiche.

---

<sup>13</sup> A description of the Early Canadian Online project technical information can be found at <http://www.canadiana.org/english/technical.htm>

	Cost	Cost per title	Cost per image	Cost per volume
Digitization	\$ 439,548	\$132.87	\$0.67	\$145.67
OCR	\$ 159,098	\$ 48.09	\$0.24	\$ 52.73
Salaries	\$ 153,264	\$ 46.33	\$0.24	\$ 50.79
Equipment & supplies	\$ 7,975	\$ 2.41	\$0.01	\$ 2.64
Construction, utilities & maintenance	\$ 21,053	\$ 6.36	\$0.03	\$ 6.98
TOTAL	\$ 780,938	\$236.08	\$1.20	\$258.82
Annual costs of storage & access	\$ 118,290	\$ 35.76	\$0.18	\$ 39.20

There are two factors that significantly lower the average cost per image of digital production and storage: the number of libraries subscribing to the database and the number of images stored. The production costs of the digital images are fixed costs that are constant regardless of the number of libraries that subscribe to the database. If there are 30 library subscriptions to the database the cost per library would be \$8.63 per volume. An increase in the number of libraries or other organizations that subscribe to the database will decrease the “cost-share” for each organization. In addition, the annual cost of storage and access to the database is also a “shared” cost. If this cost is shared among 30 libraries it decreases to \$1.31 per volume per library per year.

As the number of images available in the ECO Project increase, the cost per volume will also decline. Space costs (utilities, construction, etc.) and salaries for maintaining and updating the database and server constitute 97 percent of the costs of storage and access. These costs are incurred regardless of the number of images. Storage costs per volume are \$0.90 of annual costs. As the number of images in the database increase, total storage costs will increase, but the average cost will continue to decline.

The cost estimates from table 4 can be compared to similar recent studies estimating the cost of digital production. Estimates from studies at Cornell University and Yale University are shown in table 5.<sup>14</sup> (Cost estimates from Cornell and Yale are shown in Canadian dollars for comparison. Cost per volume is based on a 216 page text.)

	Cost per Image	Cost per Volume	
Early Canadiana Online	\$ 1.20	\$258.82	Average cost estimate
Yale	\$ 0.40	\$ 83.96	Marginal cost estimate
Cornell	\$ 0.43	\$ 91.37	Marginal cost estimate

These earlier studies show a significantly lower cost of digitization. The Cornell study created digital copies from paper while the study at Yale created digital copy from microfiche. The major difference between the Early Canadiana Online Project and these

<sup>14</sup> <http://www.library.cornell.edu/preservation/pub.htm>

earlier studies is the method used for estimating costs. Both the Yale and Cornell studies estimated costs by timing staff scanning pages of print or microfiche. These studies are based on the marginal cost of scanning images and producing digital copy. The cost estimates for the ECO project are average costs based on dividing total project costs by the number of images, titles, or volumes. The ECO Project cost analysis includes the full cost of producing digital copies and mounting the database on a server for access over the Internet. The ECO project is larger in scope, number of titles, and number of images. ECO costs include all salaries, space costs, and outsourcing of digitization and OCR. Therefore this cost analysis should be viewed as a liberal cost estimate of a large digitization project with Internet access to the database.

### *Economies of Scale*

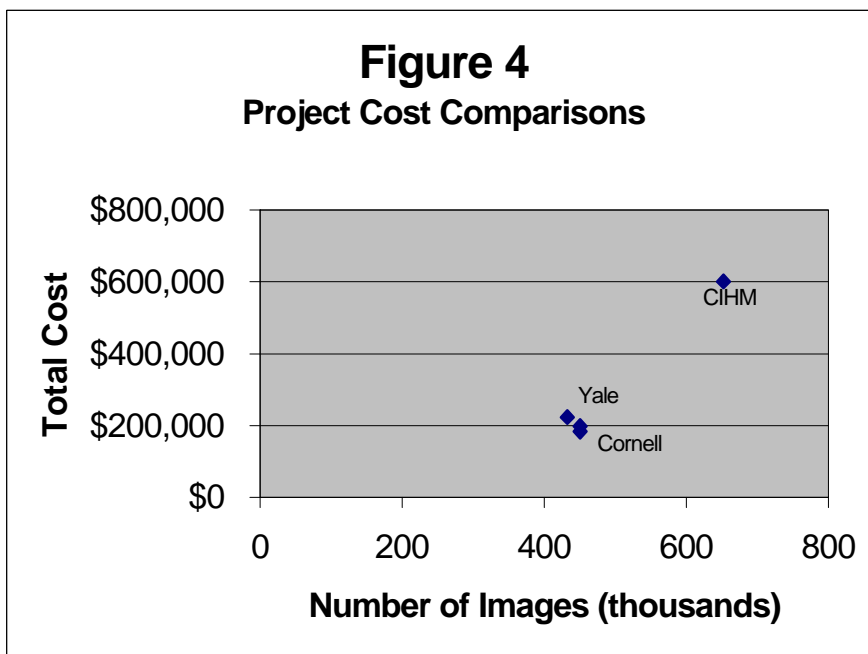
The ECO project scanned a larger number of titles and images than the projects at Yale University and Cornell University. The ECO project scanned 3308 titles compared to the 1270 titles scanned at Cornell or the 2000 titles scanned at Yale. Table 6 compares fixed, variable, and total cost estimates for the three projects.

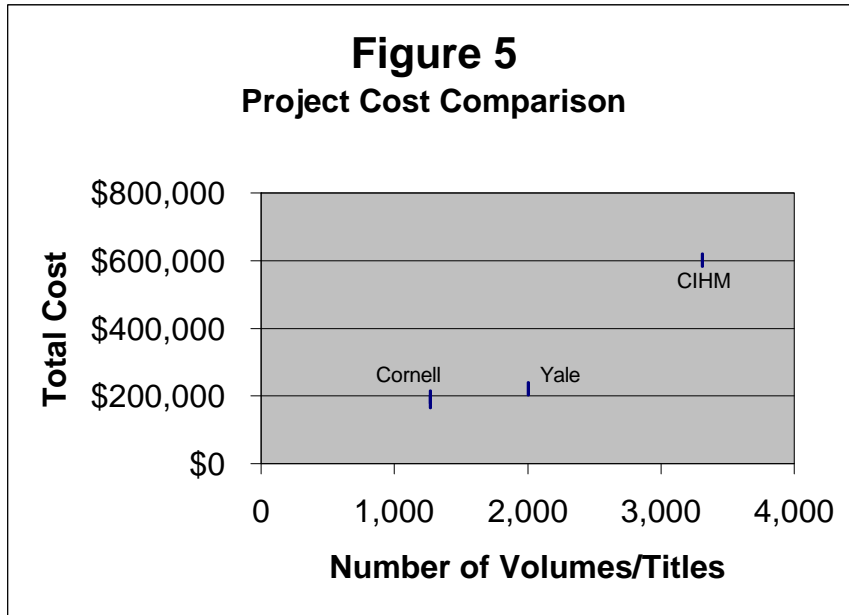
The variable cost estimates in table 6 for the ECO project include only the cost of scanning the images. OCR, space, and other salary costs contribute to total costs. For comparison with the Yale and Cornell studies, however, the vendor's cost of providing digital access may be more relevant. If more texts are digitized without OCR then the additional cost would be \$0.674 per page. The relative costs and size of the three projects are shown in table 7 and figures 4 and 5.

Table 6 Digital Project Cost Comparisons					
	Annual fixed project costs (equipment & salaries)	Per image variable cost estimates	Total Cost Estimate	Extras	Project size
Yale University Project Open Book (film to digital, 1994/95, marginal cost estimate)	\$142,420	\$0.182 (based on 600 volumes timed)	\$221,177		432,000 images 2,000 volumes
Cornell University (paper to digital to COM, 1994/96, marginal cost estimate)	\$27,931	\$0.319 manual \$0.288 auto (150 volumes timed)	\$197,275 manual \$183,353 auto	\$80,417 (COM costs)	450,000 images 1270 volumes
Early Canadiana Online (fiche to digital, 1998/99, average cost estimate)	\$161,239	\$0.674 (digitization contract)	\$600,787	\$159,098 (OCR) \$21,053 (space)	651,742 images 3308 titles

Table 7 Digital Project Average Cost Comparisons				
	Number of images	Cost per image	Number of titles	Cost per title
Yale University Project Open Book (fiche to digital, 1994/95, marginal cost est.)	432,000	\$0.51	2,000	\$111
Cornell University (paper to digital to COM, 1994/96, marginal cost estimate)	450,000	\$0.42	1,270	\$155
CIHM Early Canadiana Online (fiche to digital, 1998/99, average cost estimate)	651,742	\$0.92	3,308	\$182

Figures 4 and 5 illustrate the increase in cost per image and cost per title between the three projects. This may show diseconomies of scale, i.e. an increasing average cost as output increases. Larger projects may require more staff or have a greater complexity of task that results in higher costs per unit. However, much of the difference shown may simply be the result of different methods of estimating costs.





*Cost of Access to Digital Information*

The cost of access to digital information is difficult to quantify. Access to digital information includes the personal computer, network connection, and space used by the patron. Since these are all fixed costs of access that a patron or library must incur regardless of what information is accessed, the marginal cost of accessing any image or database is zero.

We can attempt to quantify the average cost per use to the library of providing access to digital information. This is shown in table 8.

	Cost	Cost per internal use	Cost per use
construction, utilities & maintenance	\$ 186,274.25	\$0.01	\$0.00
salaries	\$ 114,000.00	\$0.01	\$0.00
equipment & supplies	\$ 398,000.00	\$0.03	\$0.01
<b>TOTAL</b>	<b>\$ 698,274.25</b>	<b>\$0.04</b>	<b>\$0.01</b>

Table 8 includes the cost of computers within the library, staff to maintain the server and network, and the cost of space for each computer. Cost per use is shown in terms of internal use and all uses of library databases regardless of the source. Internal use is defined as the number of unique and significant hits to the library server which originate from within the library (0.3 million per week). Use is the number of hits regardless of source (1.2 million per week). Regardless of which definition of use is applied, access to

digital documents comes at a very low average cost per use. This is significantly lower than the average cost per use for microfiche or rare books.

Table 8 also illustrates the importance of understanding the difference between total, average and marginal costs. Table 8, like previous tables, shows the total and average costs per use. The total cost of providing electronic access within a university library is significant, but the high level of use of terminals within the library result in a very low average cost per use. The marginal or additional cost for each patron's use is zero. All costs in table 8 are fixed costs, incurred regardless of whether a patron uses a terminal or not. Investments in information technology within university libraries can be expensive although digital documents in a networked environment come at a zero marginal cost of distribution.

### *User Costs of Access*

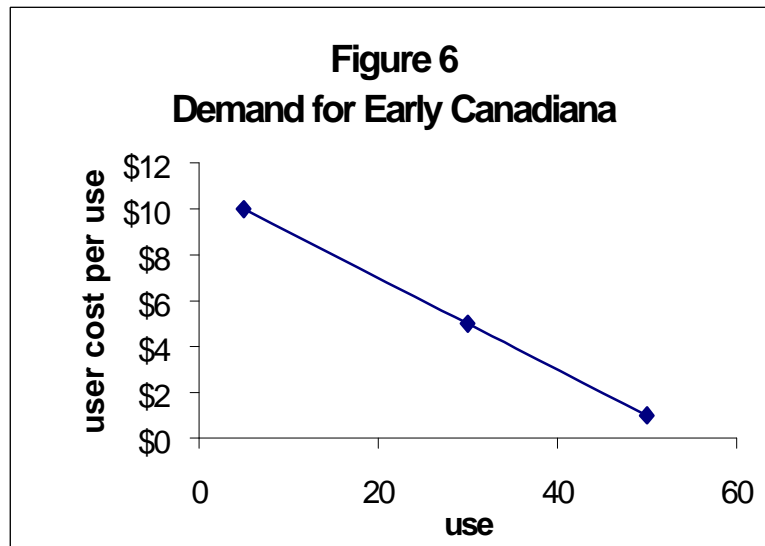
The final economic cost of access is the cost to the user. With print and microfiche the user must travel to the library to use the information. Any library will have only a limited collection of print titles. To read other titles in print from the collection a patron may have to travel to another research library. With the CIHM microfiche collection a research library can offer patrons access to a greater number of titles than are typically available in print. Although the patron must still travel to the library to access the microfiche.

Digital copies are accessible to all patrons of subscribing libraries with a network connection. This increased accessibility of the collection to patrons may result in a greater number of subscribing libraries and greater access to the CIHM collection of materials.

The cost to patrons of using information is the opportunity cost of their time spent in acquiring and consuming it. The value of access to information by patrons is reflected in the demand for using the database. The demand for use of Early Canadiana Online is illustrated in figure 6.

In theory if the user has a cost of time of \$10 per use of a manuscript in a rare books library, he may only use the manuscript 5 times a month. If the patron's opportunity cost of time spent consuming the information decreases then use will increase.

Microfiche is easier and takes less effort to use than books in a rare book library. Microfiche delivery by library staff takes less time than retrieval of a rare book. Once a patron understands how to use a microfiche reader he can view several books with relative ease. In addition, patrons do not have to travel to another library to view early Canadiana texts if their library holds the entire CIHM collection on microfiche. If we assume that the cost to a patron of accessing an Early Canadiana text on microfiche is \$5, then patron use of the microfiche will increase to 30 times a month.



Finally, digital access lowers the opportunity cost of access to the information even further. Digital access enables patrons to view the information from their personal computer in their home or office, or from a computer terminal in the library. Instant access to a large collection of images from the CIHM collection means faster, searchable access to the images. Using figure 6, if we assume that the opportunity cost of patron access is only \$2 per access for digital images, then patron use of digital access will increase to 50 uses a month.

To patrons the time savings from digital access has two parts. First, there is the value to patrons of lower cost access to images they would have traveled to the library to view on microfiche. If a patron would have used microfiche 30 times a month at a cost of \$5 per use, and this cost declines to \$2 per use in digital form, then this patron has a \$3 lower cost of access for 30 uses, or has decreased his cost by \$90 a month. Second, there are additional uses of digital access that provide additional benefits to patrons. These additional uses can be assigned an average value of \$1.50 each, or one-half of the value of lower cost access to the first 30 uses a month. If use increases to 50, the additional 20 uses per month would provide a benefit to this patron of roughly \$30. The total value to this patron would be \$120; the \$90 in lower costs plus the additional \$30 in benefit from an increase in access.

During this study, patron use of the print, fiche, and digital collection was observed. Patrons were also asked questions about their use and travel time to the library. Annual use of the collection at the University of Toronto and Laval University increased from 2984 for print and microfiche to an estimated 7,030 uses of the digital texts. Travel time to the library for print and microfiche patrons varied from less than 30 minutes to more than one day, with 90 percent of patrons needing one hour or less. The results of this survey are discussed in more detail in a separate report.

If we assume that digital access saves print and microfiche patrons 30 minutes of travel time and that the value of this time is \$10 per hour, then the annual savings of 7,030 uses equals \$25,035.<sup>15</sup> This represents a lower-end estimate of the savings from accessing the CIHM collection online versus traveling to the library to use the microfiche or print. Some patrons are likely to save more than 30 minutes of travel time. Other patrons are likely to have an opportunity cost of time greater than \$10 per hour. Most significantly, use of the Early Canadiana Online collection is likely to increase as more scholars and students are made aware of it.

### **Pricing of Digital Information in a Networked Environment**

Support for digital information projects has come from grants, donations, voluntary efforts, subscriptions, and pay-per-view access. Online scholarly journals are supported by library and individual subscriptions from existing publishers (Project MUSE<sup>16</sup>) or the voluntary efforts of a few dedicated academics and the financial support of their institutions (New York Journal of Mathematics<sup>17</sup>). Digital databases from the popular press are supported through individual annual subscriptions (Electric Library<sup>18</sup>) while newspaper digital archives are typically supported through pay-per-view access (The Albany Times Union.<sup>19</sup>) Projects, such as the Early Canadiana Online Project, which involve the digitization of existing print information, typically receive grants and donations during a startup period and then rely on revenues from subscriptions, donations, or financial support from a government organization. JSTOR<sup>20</sup> received seed funding from the Andrew W. Mellon Foundation<sup>21</sup> but now supports itself through subscriptions. Other digitization projects such as The Making of America,<sup>22</sup> as well as the Early Canadiana Online Project, have received seed funding from foundations and sponsors. The American Memory Project<sup>23</sup> at the Library of Congress receives generous funding from the US government as well as corporate and foundation support.

While donations, foundations, and government support provide initial funding for digitization projects continued revenue streams are necessary for long term financial support. Perhaps the best example of successful long term financing is JSTOR. JSTOR digitizes, stores, and provides access to back issues of scholarly journals in several disciplines. Initially JSTOR was conceived and funded by the Andrew W. Mellon Foundation. Presently JSTOR receives funding from a two-part library subscription.

---

<sup>15</sup> The calculation is (2984 uses)(\$5 saved per use) plus (7030-2984 additional uses)(\$2.50 average savings per use) equals \$25,035.

<sup>16</sup> <http://muse.jhu.edu>

<sup>17</sup> <http://nyjm.albany.edu>

<sup>18</sup> <http://www.elibrary.com> The Electric Library provides free searching of their database of newspapers, magazines, and broadcast transcripts but requires an annual or monthly subscription to view text.

<sup>19</sup> <http://www.timesunion.com> The Albany Times Union provides access to current stories and archive searching for free but requires per article payment to view stories from their archives.

<sup>20</sup> <http://www.jstor.org>

<sup>21</sup> <http://www.mellon.org>

<sup>22</sup> <http://www.umdl.umich.edu/moa/>

<sup>23</sup> <http://memory.loc.gov>

Subscribing libraries pay a one-time Database Development Fee (DDF) and an Annual Access Fee (AAF) which depend on the size of the subscribing institution. All library patrons within a set of on-campus IP address receive unlimited and unrestricted access to the JSTOR database.

This type of subscription pricing is more likely to satisfy economic efficiency than pay-per-view or individual subscription pricing. The economic theory of libraries is that the institution pays for the development of information resources that are then provided to patrons without user fees. This enables patrons unlimited consumption of books, journals, and databases that are available in the library. Since the economic cost of lending a book, journal, or enabling a patron to view a database is lower or near zero, economic theory dictates that patrons not be charged a user fee. The fixed costs of publishing journals or developing databases are paid for through fees charged to the library which are in turn paid through tuition revenues or taxes. The difficult task of the librarian is to determine what information products to purchase for patrons with unlimited wants for access.

Using the cost of production from table 4 and assuming 30 subscribing libraries, efficient pricing would result in each of the 30 libraries paying a cost of production charge of \$7.87 per title for production and an annual charge of \$1.19 per title for storage and access. It is anticipated that the storage and access charges will decline with the decrease in computing storage costs and the increase in the number of titles to which employees are responsible for providing access.

The number of libraries that would be potentially interested in this database is significantly greater than 30. The Early Canadiana Online database is potentially of interest to research, academic, special, and school libraries. While microtext requires shipping, production, and distribution costs, access to the digital copies of these texts requires only a computer with an Internet connection. Estimates are that 90% of schools and 65% of public libraries in Canada have Internet access.<sup>24</sup>

*Potential Market Size*

The most difficult task is to set prices for access to the database. Potential revenues and market size are shown in table 9. In table 9, DDF is defined as the database development fee. AAF is the annual access fee.

---

Table 9

Revenue Estimates for Two-part Fees for Early Canadiana Online						
Total Number <sup>25</sup>	DDF	AAF	Estimated number of subscribing libraries	Revenue DDF	Revenue AAF	

<sup>24</sup> Estimates are from Canada's SchoolNet, <http://www.schoolnet.ca>

<sup>25</sup> Library estimates come from American Library Directory 1998-99, RR Bowker, New Providence, NY and Ralph Manning "Counting Libraries: Reading Between the Lines" *Focus on Culture*, 9:2 (Summer 1997).

Academic	487					
Very large		\$25,000	\$2,500			
Large		\$20,000	\$2,000	45	\$675,000	\$67,500
Medium		\$15,000	\$1,500			
Small		\$10,000	\$1,000			
Very small		\$ 5,000	\$ 500			
Public	782 (excluding branches)	\$ 1,000	\$ 200	90	\$90,000	\$18,000
Government	389	\$1,000	\$ 200	50	\$50,000	\$10,000
Special	1422	\$1,000	\$ 200	50	\$50,000	\$10,000
School	14000	\$0	\$ 100	500	\$0	\$50,000
Foreign		\$10,000	\$1,000	10	\$100,000	\$10,000
<b>TOTALS</b>					<b>\$965,000</b>	<b>\$165,500</b>

Unfortunately the revenue estimates shown in table 9 would provide only enough financial support for the database of 3308 titles currently in the database. Given the market size of libraries potentially interested in ECO images there is limited potential for revenues sufficient to digitize the entire CIHM collection.

There are three possible solutions to increasing revenues for further digitization. Firstly, the fee structure in table 9 could be changed to raise sufficient revenues. The difficulty is finding the correct fee structure. In addition, libraries currently subscribing to the microfiche collection may be unwilling to also subscribe to the digital collection. A discount for libraries that subscribe to both should be offered. This would, however, only decrease revenues provided solely for the digitization of additional microfiche. Secondly, additional financial support from foundations may be sought to increase the number of items digitized. Thirdly, corporate, government, and organization support could be sought to increase the digital collection. For example, table 9 assumes that school libraries will not be charged a DDF. Given the resources of most school libraries it is unlikely that many libraries would subscribe if this fee were charged in addition to the AAF. However, the ECO database would provide enormous value to school children through Canada and the rest of the world. It may be possible to solicit corporate, government, and foundation sponsorship for schools throughout Canada to provide additional financial support for the ECO project.

## Conclusions

The economic paradox of digital information is finding the correct financial strategy to collect sufficient revenues to pay for the benefits of digitization. Digital information provides greater access to information at a lower cost. However funding the production, archiving, and access to the information requires creative financing including value based pricing of information as well as the solicitation of grants and donations.

Information production and access comes at a cost. An accurate measurement of the full economic costs of different methods of information delivery is essential in determining the most cost-effective method. This study has shown the costs of three methods of access; print, microfiche, and digitization of microfiche. The cost of digital information is lower

on a cost per library or per patron basis so long as a sufficient number of libraries are interested in subscribing to the database.

In general, the lower cost of digital production will continue to result in more information products appearing in digital format on the Internet. The increase in the number of digital products will further contribute to the information overload of patrons and librarians. Information consumers are confronted with too many journals, databases, and research sources for the limited amount of time and attention they can give to any one source. Given a limited amount of time for information consumption, patrons will search for information of higher quality for use of their time. Any new digital product must have an assurance of quality in order to convince patrons and librarians that there is value in spending time consuming it. Manuscripts of historical significance, such as the ECO Project, produced by trusted organizations, such as CIHM, provide libraries and patrons with an assurance of quality.