A VISION FOR INNOVATION IN SCHOLARLY BOOK PUBLISHING

The Chicago Digital Distribution Center and BiblioVault

Contents

Executive Summary 1

Introduction 7

Background 9

The CDC Is an Appropriate Host for A Pilot Project in SRDP and In Developing a Digital Book Repository 13

A Timely Response To The Challenges Of Scholarly Publishing 15

CDDC and BiblioVault—Organization for First 18 Months 34

CDDC and BiblioVault—Projected Calendar Of Activities 36

Executive Summary

The University of Chicago, acting through the Press and acting as the agent for the scholarly presses in the Chicago Distribution Center (CDC) and for other interested not-for-profit scholarly publishers, is requesting from The Andrew W. Mellon Foundation a grant of $1.5 million to fund an 18-month pilot project to establish a short-run digital printing center and a digital book repository, the BiblioVault, in the newly created Chicago Digital Distribution Center (CDDC). The University of Chicago Press will manage the CDDC as an offshoot of the CDC, but it will, of course, consult with consortium members about policies, procedures, problems, and plans for future activities.

The economics of scholarly publishing, never attractive given the range of intermediaries, high discounts to wholesalers and retailers, and a policy of unlimited returns, have grown increasingly difficult in recent years. In addition, scholarly publishers see value in adopting new technologies for the digital production and delivery of books, but we lack the funds to make the needed investments. This grant would enable a significant set of not-for-profit scholarly presses to establish a firm foundation for moving forward in this new century to a sounder financial footing for publishing new books and reprinting existing titles.

Without such support, these university-based publishers will operate at a significant disadvantage in relation to larger publishers committed to making serious nonfiction broadly available using state-of-the art production technologies. Such innovation on the part of commercial firms will enable them, increasingly, to "cherry-pick" work coming from the academy. That is they will target those manuscripts that are likely to yield a profit. In order to recover our costs and to continue to identify and pub-

lish scholarship, university presses must be in a position to publish the more significant scholarly work and to make it available in multiple formats. Without continuing contributions from these better selling works, scholarly presses will not be financially able to support the publication of works that have a small academic audience and cannot pay for themselves.

Without such support, university-based publishers will continue to work independently and without benefit of technical expertise. We will be forced to respond in piecemeal fashion to the growing number of commercial firms that offer a subset of services related to the production in short runs of printed books and to the design and dissemination over the World Wide Web of electronic books. Over the past year, press directors and production managers have been eager to integrate short-run digital printing (SRDP) into our activities, but confused by the range of "deals" offered and reluctant to yield control over the creation of electronic files of our books. The approaches made by representatives of profit-seeking entities have been inconsistent with respect to the terms on which university press publishers could retrieve usable copies of the files for our own books. Firms which have undertaken and financed file conversion in order to launch electronic book products typically refuse to share these files with the publisher of origin; in fact, they often demand the right to use those files in any way they choose in the future.

This document proposes an initiative that has major implications for the ability of scholarly presses to act together in realizing the economies of scale essential to the efficient production and dissemination of books with a relatively small but important global audience of researchers and readers. Barring a coordinated effort, these university presses will not have the wherewithal to create and manage a nearly automatic system to manage the reprinting of books in small quantities; nor will we have the funds necessary to invest appropriately in book file preparation and in the systems that will allow us to fully integrate short-run digital printing into our publishing operations.

Scholarly presses recognize that managing electronic files for future needs will be critical to our financial well-being and to our ability to serve the scholarly community fully. Again, individual presses do not possess the expertise or resources to plan for very different modes of producing, selling, and delivering books. Only by working and learning as a group, with the backing of a visionary organization like The Andrew W. Mellon Foundation, can scholarly presses take the steps that will ensure our con-

tinued viability in times marked by changing forms of scholarship, distribution of scholarly works, and technology.

In a Fall 2000 survey, we found that scholarly presses are grappling with the challenge of how best to manage low selling titles, inventories, and decisions on reprinting titles with modest sales. Indeed, many university press titles have sales of fewer than 1,000 copies over their lifetime. Seeking to achieve the lower unit costs of large print runs, university presses tend to print more copies than are ever sold. For universities and their publishing arms, the result is over-investment in inventory, funds wasted on warehousing, and ultimately the pulping of excess inventory. If university presses reduce over-production, we may use the funds recovered to publish more books and otherwise to further the goals of scholarly publishing, such as achieving greater dissemination of books that we publish—new titles as well as backlist books that we cannot afford to promote on a continuing basis.

American university presses typically operate book-publishing programs at break-even or at a loss — usually subsidized by their host institutions — and we lack funds to invest in preparing our books to use new production technologies and in developing management systems to use these new technologies efficiently. SRDP is a relatively new technology that allows publishers to produce small runs of books (25 to 300 units), as they are needed and to maintain minimal inventory. The unit costs for SRDP are higher than those for standard offset printing, but SRDP allows a publisher to eliminate its risk of over-investment in inventory. As the quality associated with this technology has improved dramatically in the last year, scholarly publishers would like to be able to adopt SRDP now. The grant proposed herein would allow CDC presses to improve our management of our titles, enhance our finances, and make individual titles quickly and easily available through the short-run digital printing facility at the Chicago Digital Distribution Center.

As the scholarly literature grows, scholars find it increasingly difficult to locate and obtain the books and journal articles that are relevant to their work. The CDDC will include the digital book production center at the CDC and a repository (the BiblioVault) for the digital files needed to produce these books. The BiblioVault design will enable not-for-profit scholarly presses to maintain control over our intellectual property — the books we have published — and to deploy it for multiple purposes around the world as the use of digital files for books expands. At this time, commercial firms

are pressuring us to give them our published material for use in enterprises created to capture revenues from the individual scholar and library markets.[1] The longer-range goal of the BiblioVault is to provide a repository for scholarly presses that will serve the scholarly community by enabling presses to provide our books to libraries and scholars in print and electronic formats at prices and in packages (of one or many books) that serve the community by both making the material available and maintaining financial health for libraries and for university-based presses. Scholarly presses, libraries, and readers will benefit, as our goal is to nurture scholars and to disseminate their work broadly, while the goal of these commercial firms is to maximize profits. Retaining control of this scholarship within the scholarly community will ensure that (1) our online libraries will be designed as our libraries and scholars desire; and (2) any earnings from these online libraries will be plowed back into the creation and distribution of scholarly books and journals, just as the earnings from print works are.

While this proposal requests funds to establish the CDDC and BiblioVault over the next 18 months, this request should be viewed in the context of an effort that will unfold over the next several years. If the scholarly publishing community comes together as envisioned, this broad initiative will include:

1. An ongoing consortium of not-for-profit scholarly presses. At present, nearly twenty presses use the warehousing, fulfillment, and business services of the CDC.
2. The Chicago Digital Distribution Center (CDDC), with (a) the digital book center to print small quantities of books; (b) a management database to highlight initially which books are appropriate for preparation for SRDP based on their sales and inventory levels, and later the date when each such book needs to be printed and available (particularly important for classroom texts), and the number of copies to print;[2] (c) the CDC warehouse to retain a small number of copies beyond those needed for immediate orders; (d) the CDC fulfillment system to dispatch orders to buyers; and (e) the CDC business office to work with consortium members to set rules and to learn about successful behaviors. The CDDC's SRDP operations will be fully developed during this grant period. This proposal requests funds to support the start-

1. Questia, ebrary, netLibrary, and Ingram are the firms with the greatest name recognition, but dozens of commercial firms have approached university presses to request rights to use their books in various online books ventures. These firms offer no upfront payments to publishers. Some require a press to provide book files in the form they will use. NetLibrary sells individual titles to libraries, typically at hardback prices, and returns 50 percent of that price to the publisher. The other firms use a variety of formulas for payments to presses, but the presses together typically receive about 15 percent of the gross revenues from these electronic book collections.

2. The publisher of a book will determine whether it wants to include it in the SRDP program, based on physical characteristics or other issues. These are modifications of the database for the current CDC operations.

up of the CDDC; areas of support include staff capable of managing new technical initiatives, the modification of a major inventory/fulfillment database, and preparation of various instructions and management materials for participating presses.

3. The BiblioVault repository to hold digital files for books, including (a) scanned files of the pages of older books for use in SRDP, and ultimately for viewing online; (b) text files of the pages of backlist titles, to be created via optical character recognition treatment of the scanned images, for use in online searching and delivery of search results; (c) electronic files of recently published books and books about to be published, for use in SRDP when the initial inventories printed by offset are depleted, and ultimately, for use in searching and viewing online; and (d) metadata catalog databases encompassing information about all of these books, such as the ownership status, or rights to various components of the books, and other metadata about the books.

   This proposal requests funds for scanning 2,300 books published in the past (backlist) and for preparing 2,850 recent books with electronic files for use in SRDP. For backlist titles, we will also create the text files and metadata. The files of the recent books will also be prepared for online use at modest cost. The proposal also requests funds for the design and creation of the databases to manage the repository and for the hardware and software components of the BiblioVault. Finally, it requests funds for a BiblioVault general manager and the technical team to launch and maintain the repository.

4. BiblioVault "products" that include collections of online books, both current and backlist. In order to be able to offer such collections to the scholarly community, delivery vehicles and related software are necessary. To build the overall infrastructure of the BiblioVault, this proposal requests funds to develop prototypes of such delivery vehicles and software tools. However, our intention is to make such collections available for sale only after the term of this proposal.

For nearly three years, university presses have been targets of commercial vendors seeking "content" that can be manipulated and sold in print and electronic form to scholarly institutions and individuals. We are naturally skeptical of the plans of such commercial vendors to use the scholarship we have published, usually with substantial subsidy from our universities, or using the surplus from the small number of titles that do more than break even. The presence of new profit-seeking entities in the chain of intermediaries from author to reader has the potential to make access to the work of scholars even more costly than it is today. The CDDC and the BiblioVault will ensure access to scholarly materials at reasonable prices and will allow scholarly presses and authors to benefit from whatever financial contribution their publications generate.

Introduction

This proposal to The Andrew W. Mellon Foundation requests a grant in the amount of $1.5 million for an 18-month project to establish the Chicago Digital Distribution Center (CDDC), its digital book production center for short-run digital printing, and the BiblioVault book repository under the umbrella of the Chicago Distribution Center (CDC). It is predicated on the desire to improve the economics of not-for-profit scholarly publishing, to keep decisions affecting access to scholarly publications in the hands of university-based publishers, libraries, and scholars, and to preserve and make available the work of scholars in its original form for posterity. short-run digital printing affords new efficiencies in the production of scholarly books to the benefit of scholars as authors, researchers, instructors, and students. This project will also lay the groundwork for online access to scholarly publications directly from academic institutions.

While this project is invaluable to the community of scholarly presses, individual presses do not have the resources to undertake such a cohesive set of ventures alone; nor do we possess the infrastructure and funds necessary to act as a group. Universities subsidize most of these publishing organizations; our book programs operate at break-even or at a loss. With modest endowments or reserves, we cannot fund technological initiatives on behalf of our broad community of not-for-profit book publishers. And perhaps most compellingly—and urgently—no single American university press enjoys the economies of scale that commercial publishers and printers have used over the past several years to justify heavy investment in digital technologies. The vision behind the CDDC is the creation of a major high technology book production center married to an efficient distribution and fulfillment center that brings just such economies

of scale to the scholarly publishing community. Individual university presses will be able to focus their expertise on the editorial and publicity functions essential to the publication of scholarship that reaches a global audience in an era when the resources available to university presses are increasingly limited.

Funding from The Andrew W. Mellon Foundation would enable an important group of not-for-profit publishers, many based at major research institutions, to launch the Chicago Digital Distribution Center and BiblioVault, and to demonstrate that production of books in small quantities has the potential to allow books to stay in print for all time. Enhanced cash flow from the elimination of excess inventory and from the sale of books previously out of print should enable participating presses to contribute to the expansion of BiblioVault activities in later years.

Background

THE UNIVERSITY OF CHICAGO PRESS

The University of Chicago Press (UCP), an integral part of the vision advanced by William Rainey Harper for a university with a publications arm dedicated to the dissemination of faculty research, was founded in 1891. The publishing program at the University has always supported both books and journals. A third division, the Chicago Distribution Center (CDC), devoted to warehousing, customer service, and order fulfillment for books, was added to the operations of the Press in 1965. The CDC now provides warehousing, fulfillment, and business services for nearly twenty not-for-profit publishers.

Since the arrival of Paula Duffy as director of The University of Chicago Press in August 2000, the organization has been re-structured with greater coordination of the technical, business, and Web-based marketing areas of the Press. Mary Summerfield, director of business development and planning for UCP, was formerly the project director at Columbia University who had primary responsibility for the Foundation-funded Columbia Online Books Evaluation Project. She will direct this project. Evan Owens, formerly electronic publishing manager for the Journals Division, is now information technology manager for UCP. He will oversee the information technology aspects of this project. Lain Adkins, formerly associate director for finance and operations at the University of Oklahoma Press, is UCP associate director for operations. He will work closely with Don Collins in managing the CDDC.

GOALS OF UNIVERSITY PRESSES IN THE CONTEXT OF DISMAL PUBLISHING ECONOMICS

KEEPING BOOKS IN PRINT A core goal of university presses is to maintain in print and at reasonable prices the books that libraries and scholars seek to purchase. The University of Chicago Press (UCP) has about 2,500 paperback titles in print and over three million paperback units on hand. UCP rate-of-sales and inventory issues are typical for scholarly presses.

Low and irregular demand characterize the sales of university press backlist titles. About 28 percent of UCP backlist paperback titles sell 100 or fewer units a year; 74 percent of UCP's paperback titles that have had only one printing sell 100 or fewer units each year. While the UCP strategy is to print for three years' demand, we cannot estimate demand precisely and we seldom meet that goal. Some books run out of stock before three years, but more often demand is over-estimated and books linger in the warehouse.

The capital invested in book inventory is substantial. This is particularly troublesome for the roughly 50 percent of university presses that operate on a cash basis. Funds invested in inventory are not available for other press activities such as developing books. Presses could redirect funds to efforts of value to the scholarly community if they could meet the demand for their books in a more cost-efficient way.

Book publishing has been a fundamentally difficult undertaking since the 1930s when conventions arose that allowed wholesalers and retailers to return books to the publisher for full refund at any time. The incentives for wholesalers and retailers to manage their operations efficiently have been small and publishers of scholarly monographs have operated with slim to negative margins.

In recent years, scholarly book publishing has become even more difficult. The return of unsold books is at an all-time high as wholesalers and retailers go out of business, close warehouses, and change their ways of conducting business. Scholarly presses are hard pressed to maintain sales volume as the major retail chains and wholesalers reduce their advance orders for new titles. An increasing amount of wasteful activity characterizes the scholarly publishing industry, e.g., books are returned shortly after they are published and then reordered and restocked after reviews and commentary are published.

In addition, it is becoming increasingly difficult for publishers to establish titles as part of their ongoing backlist. Traditionally, books published in preceding years have accounted for about two-thirds of university press

revenues.[3] The revenue from the backlist is shrinking, however, as courses adopt fewer titles in book form and students are less inclined to purchase assigned books. Scholarly presses need greater awareness of the patterns of purchase and use of scholarly books. And we need to adopt new strategies that will ensure our financial viability.

Short-run digital printing (SRDP), in which a few copies of a paperback book can be produced quickly and relatively economically, is a solution to the challenge presented by the goal of keeping books in print; it also allows initial publication of some titles that would not be economical with traditional printing technology. Digital printing technologies should allow university presses to recover our investment in books while meeting annual demand for units in the tens or low hundreds, a level which precludes economical offset printing and which traditionally has led to books going out of print.

HELPING SCHOLARS IDENTIFY AND OBTAIN USEFUL TEXTS Another important goal of university presses is to provide scholars with full intellectual and physical access to the knowledge contained within the books we publish. At present, however, such access is inadequate. Library catalog records say little about a book, so research scholars, instructors, and students miss books that may be relevant to their work even when their library has them. The popularity of online bookstores within the academic community reflects the high value scholars put on the ability to search online databases and identify work by a specific author or on a specific topic. In addition, the exploding use of online full text databases of journals, such as JSTOR, reflects scholars' increasing comfort with online research.

Physical access to scholarly literature has always presented problems for researchers because most library collections contain only a small portion of the literature related to a topic. Multiple factors make it difficult for libraries to acquire all the books their patrons might want to use. Budgets strained by the drastically increased cost of journals (particularly scientific, technical, and medical journals from commercial publishers) are one key factor. The volume of scholarly publishing is so large that few libraries can afford to acquire, prepare for the shelf, and house

3. Sixty-three presses reported backlist shares of revenues of 62 percent to 71 percent for 1998 to the AAUP.

more than a modest subset of the books published each year.[4] Acquiring a book and getting it to the stacks costs over $40, thus potentially doubling the cost of a book to a library.[5] Maintaining a print collection is also costly.

Even if the library owns a book, a scholar often cannot obtain it when it is most needed. Books are unavailable because they are circulating, waiting to be reshelved, shelved incorrectly, lost, or set aside in the reserve collection. In general, for both libraries and their scholar patrons identifying and obtaining relevant paper books is difficult, costly, and time-consuming.

The core business proposition of commercial online book and library vendors, such as netLibrary, Questia, and ebrary, reflects the recognition that print volumes are often not accessible when and where a scholar or student may want to use them. Libraries have limited hours, and scholars who live away from campus or are traveling have needs that currently go unmet. Interlibrary loan is slow for the user and costly for both the borrowing and lending libraries.[6] Online collections of scholarly books and journals are a partial solution to these problems for college and university libraries and scholars. The more scholarly works contained in such collections, the more valuable they will be to scholars. Of course, some scholars will continue to want to read entire books and will require printed copies. But the amount of use that most scholars make of most works is such that an online copy and limited printing or copying will suffice. And for an initial review of a book, immediate online access is certainly preferable to delayed access to a print copy.

In later phases of the BiblioVault, we will work to make the backlist and new books of the participating scholarly presses available for full-text searching and reading online. In the latter months of this first phase, we will test search capabilities with prototypes for online electronic books, clearance of rights for such books, and delivery systems including search engines and Web pages.

4. The Association of Research Libraries recently issued 1999-2000 statistics for 112 members. Harvard added the most volumes (289,322) and the University of Guelph added the fewest (26,007). Four libraries added more than 200,000 volumes; 27 added 100,000 to 200,000 volumes; 59 added 50,000 to 100,000 volumes; and 22 added 25,000 to 50,000 volumes. The mean number of volumes added by these research libraries was 85,795. These are the largest universities; most colleges and universities acquire far fewer volumes (books and journals) each year.

5. The final report of the Columbia Online Books Evaluation Project contains an analysis of a library's lifecycle costs of owning print and online books on page 142. It is available at <http://www.columbia.edu/cu/libraries/digital/olb-docs/finalreport.pdf>.

6. Interlibrary loan is a costly and inadequate substitute for owning a book. Its full cost is close to the price of a book; a scholar waits weeks to receive the requested book, if he receives it at all, and may then find that it is not useful to his work.

The CDC Is an Appropriate Host for A Pilot Project in SRDP and In Developing a Digital Book Repository

Conducting short-run digital printing at the CDC, where the books are stored and shipped, will eliminate the need for client presses to send SRDP books from the printer to the warehouse.[7] The CDC database maintains information on stock levels for client presses. CDC programmers who are experts on this database will modify it so that the critical decisions on the need for a print run and the quantity of books to be printed are nearly automatic. The resultant savings in management time will allow press directors and their marketing and production staff to focus on more significant publishing concerns while the books themselves remain available for scholarly research and classroom adoption.

The CDDC manager and production coordinator will discuss and provide data and advice on best practice related to short-run digital printing and electronic file management that will help participating presses to understand the opportunities presented by SRDP. In addition, the consulting technical writer will prepare various instruction sets and database manuals that will enable these presses both to provide their materials to the CDDC in appropriate form and to extract information from the management databases in ways that will enable them to make better business decisions.

THE JOURNALS DIVISION EXPERIENCE WITH ELECTRONIC PUBLISHING WILL FACILITATE DEVELOPMENT OF THE BIBLIOVAULT

Our initial work in the design and operation of a repository for book files, the BiblioVault, will benefit greatly from the experience of the Journals Division in developing the ability to review, edit, and deliver journal arti-

7. It is likely that in the future, publishers will also be able to have their books printed at sites convenient to the scholars who wish to purchase them, e.g., at bookstores or print shops. This would have a dramatic, positive effect on the rate of book returns. The BiblioVault will be designed to deliver book files to such distributed printing facilities around the world.

cles online.[8] The expertise of Journals information technology managers and electronic publishing staff will enable us to design and build the BiblioVault repository, including its storage system and technical catalog, and ultimately to provide searchable collections of electronic books. Internal information technology and electronic publishing staff, hardware, and software resources will be assigned to the BiblioVault initiative as needed and are reflected in the project budget.

8. The UCP Books Division has limited but valuable experience in designing and launching electronic products as well. Among these are *The Founders' Constitution* and *The Chicago Homer.*

A Timely Response To The Challenges Of Scholarly Publishing

OUR VISION OF THE CHICAGO DIGITAL DISTRIBUTION CENTER AND BIBLIOVAULT

The CDDC will be an offshoot of the CDC. From Fall 2001 through Winter 2003, the CDDC will (1) develop and provide SRDP services to not-for-profit scholarly presses; (2) initiate the participation of university presses, primarily those who currently participate in the CDC and those who intend to join the CDC; (3) process 2,300 print books with low demand that might go out of stock during this period so that they can be reprinted via SRDP as needed and so that, in the next phase of this effort, they can be provided to libraries and scholars electronically; (4) process the electronic files for 2,850 books recently published or about to be published so that they can be used for SRDP when appropriate and so that, in the next phase of this effort, they can be provided to libraries and scholars electronically; and (5) design, build, and maintain a robust repository, the BiblioVault, which will contain, manage, and deploy as needed the various forms of digital files for just over 5,000 books initially and many more ultimately.

Participating presses will use the CDDC's digital book production center to print small runs (25 to 300 copies) of books with modest or uncertain sales potential as needed to serve the scholarly community. CDDC managers will work with these presses to develop a better understanding of the economics of book production, to rethink production strategies, and to reduce inventory investment. While backlist books that have small sales and whose stock is depleted are the most likely candidates for SRDP, this technology is also appropriate for advance copies of new books, reprints of recent books, or even initial printing of books with unknown or small demand.[9]

9. New or backlist books can also be printed in hardcover format using SRDP to create the pages and standard binding technology for the cover.

SRDP: A PARTIAL SOLUTION FOR THE BOOK PRODUCTION QUANDARY

As we noted earlier, a principal goal of not-for-profit scholarly publishers is to keep our books in print. The purchase pattern for many older titles suggests that virtually all of their sales result from their inclusion on required reading lists for one or two courses, often one course every second year. Such demand results in the sale of perhaps 25 to 50 copies when a particular class is offered. If possible, university presses do provide the needed copies. However, once inventory is depleted, the publisher needs to decide whether to go back into print or to let the book go out of print. Publishers have developed various rules for this decision, but the decision for each title typically demands management attention, as it is not automated. Short-run digital printing is a solution to this quandary for this class of books.

SRDP has been available for several years, but the quality of printing of pages and covers initially was poor compared to standard offset printing. Quality of printing—both texts and covers—has improved dramatically over the past year and continues to improve. Offset quality paper is now used and covers can be four color and laminated. However, halftone photo images printed with SRDP are still inferior to those printed conventionally.

SRDP has many benefits for scholarly presses and the scholars they serve. To summarize, SRDP:

- Minimizes the publisher's inventory investment, thus releasing cash for other uses and reducing the risk involved in the printing.
- Offers potential as the printing method for the initial printing of some monographs that are expected to have very limited sales.
- Prompts reconsideration of the status of books that have long been out of print and for which there appears to be limited demand.
- Allows publishers to keep newer books in print and available in perpetuity.
- Allows presses to print their own advance copies of books for distribution to the book review media (a traditionally high-cost activity essential to the promotion of books with a potentially broad general audience).
- Addresses an area of need already identified in the market place, as our survey of scholarly presses showed.
- Makes available to scholars what they want in a form they want. (Even if books are available online, some scholars want them in printed form.)

THE ECONOMICS OF PRINTING

SCHOLARLY BOOKS

Scholarly publishing is characterized by many overhead and fixed costs—at one level in the production of a list of books, at another level in the production of a single title, and finally in the production of the copies of a book.[10] Once a book has passed the stages of acquisition and developmental editing, copyediting, production editing, and design, it goes on to typesetting, and finally to printing.

In the initial offset printing of a book, the publisher incurs fixed costs of proofs, preparation of the film for the base text and cover, printing plates for the base text and cover, preparation of the press for printing the text and cover, and finally preparation of the binding machine. The costs of paper, printing, and binding (PPB) are variable with the number of units produced. As each preparatory step is necessary to produce books (no matter how few copies), the fixed costs of a print run are high and the marginal cost per unit is low and declining as order size increases.

We have analyzed the costs associated with producing original and reprint paper editions using offset printing technology and costs of an SRDP reprint. This analysis is for a book with the following specifications: paperback, 6" x 9" trim size, text-only, black ink, 248 pages, 50 pound natural paper, perfect binding with a four color cover. For an original printing of 5,000 copies — a relatively large order for a university press, but a small one in the commercial world — the total cost per unit is $1.59. For an original printing run of 750 copies—more copies than many university press titles ever sell—the total cost per unit is $4.39.

Reprinting a book via offset eliminates only the costs of proofs and preparation of the film; new plates must be prepared for each printing. Thus, the costs are only moderately lower than those for the first printing, $1.36 per copy for a run of 5,000 copies and $3.10 per copy for a run of 750 copies.

These high fixed costs and low marginal costs tempt the publisher to print more copies than needed when submitting orders for new titles or reprinting their out-of-stock titles. In our example of an original printing, doubling the print run from 750 copies to 1,500 copies increases total costs by just 24 percent. This equates to increasing the investment in inventory by $789. If that is the amount by which the print run exceeds demand, and if a scholarly press prints 100 titles a year and typically over-prints by this

10. For a brief general discussion of these costs, see *Issues in the Economics of Scholarly Communication*, by Mary Summerfield (revised March 1998) <http://www.columbia.edu/cu/libraries/digital/texts/econpap.html>.

amount, the press will have over-invested in inventory by $114,900 that year.[11] Such a press is likely to have annual revenues in the vicinity of $3 - $5 million, so this over-investment equates to about 3 percent of revenues. As university presses are typically subsidized to this level,[12] this is a meaningful sum that could be put to good use furthering the dissemination of scholarship.

SRDP has a very different cost structure from offset printing. Its one time preparation costs vary substantially depending on whether a book must be scanned or whether the press has a PDF file available. The fixed costs of this process are preparing a file of the pages and cover for SRDP and storing that file over time. SRDP printers typically charge about $.25 per page for scanning basic pages.[13] Scanning images to a higher quality level is more costly. Preparing a cover file from a cover or jacket is labor intensive and charges are typically around $50 per cover. Preparing a proof copy of the book costs $10. Archiving the files costs $25. Thus, the total mastering charge for a print book for which a PDF file is not available costs about $150 for a 248-page book. Preparing an SRDP file of a book's pages from a PDF file typically costs about $25. Thus, the capital costs of SRDP for books for which PDF files are on hand (i.e., those produced in recent months) will be much lower than costs of file preparation for books that must be scanned.

The printing cost analysis finds that the total cost of an order for 300 copies is roughly the same for SRDP and for offset printing if one ignores the fixed cost of preparing the SRDP file. However, printers take offset orders for 300 copies only reluctantly and typically do not accept orders of fewer than 500 copies. This inhibits the ability of university presses to keep books in print for which a continuing, but small course adoption market exists. The actual costs of producing a copy of a book by SRDP declines slightly as the number of copies in an order grows; it is just over $6 with the system that will be in place at the CDDC. The more copies over which the mastering cost is amortized over time, the lower the total cost per copy. If a single copy carries that cost, the total cost of that copy is about $153.

11. This includes an estimate of the cost of maintaining these excess copies in inventory for just two years as $0.24 per copy per year (the CDC's cost to warehouse a single copy).

12. In 1998, the last year for which these statistics were gathered, 63 university presses reported to the AAUP a total of $6.3 million in cash subsidies, for an average of $100,000 per press. Presses also receive subsidies in the form of free space and other non-cash items.

13. Having large quantities of books scanned by other vendors can be much less costly. We propose to work with the Digital Library Production Service of the University of Michigan Library. DLPS will scan, OCR, and create metadata for our books for $0.30 per page.

On the other hand, if that cost is amortized over six print runs of 50 copies, the cost per copy is less than $7. The per-unit cost of an offset reprint of 300 copies is just 14 percent lower.

The decision on whether to use SRDP or offset printing technology must take several factors into account: availability of cash to fund book inventory, the cost of warehouse space,[14] the anticipated level of sales, and the certainty of those sales. Offset printing is considerably less expensive than SRDP at levels above 300 units, but if the books do not sell, their real cost will be higher. In general, a press maximizes its production flexibility and minimizes its financial risks if it uses SRDP for books that are expected to sell about 100 copies a year or fewer.

SRDP VS. POD There is no absolute distinction between SRDP and print on demand (POD). Print on demand (POD) is often used to describe the production of the quantity of books needed to fill a single order or the orders that arrive over a designated period, e.g., a day or a week. Such POD titles would not be placed in inventory and would, in fact, be deleted from records showing warehouse stock.

SRDP often implies the production of small quantities (25 to 300) to satisfy both immediate demand and short-term inventory requirements. Because each printing involves some small set-up and overhead costs and, when SRDP production occurs at printing plants, shipping costs, publishers with available warehouse space have typically favored SRDP over strict POD production. The dynamics of this distinction change when the SRDP facility is located near the publisher's shipping facility and when SRDP production is increasingly automated through programming business rules into the inventory control/fulfillment system, as will be the case at the CDDC. Early SRDP experience and POD experiments will provide data to inform our decisions about the minimum efficient levels for SRDP and the wisdom of instituting POD under specific sets of circumstances. The distinction between SRDP and POD may collapse, particularly for titles with erratic and unpredictable sales patterns.

SRDP AT THE CDDC SRDP is typically conducted at printing plants, and books are then shipped to a warehouse that serves the publisher. Edwards Brothers, a printer based in Ann Arbor, Michigan, that has served

14. Many universities that formerly "gave" or donated space for warehousing have changed their policies in recent years. The UCP cost of warehousing a book is about two cents per month.

scholarly presses, including the University of Chicago Press, for over 100 years, is innovating with a program to locate SRDP production in publishers' distribution centers. We are pleased to work with Edwards Brothers on such a program because (1) they have an outstanding reputation in our industry, and (2) we have been very satisfied with the substantial amount of SRDP work that they have done for UCP over the past two years.

We are negotiating with Edwards Brothers to establish a digital book production center at the CDC as a core component of the CDDC. At least initially, Edwards Brothers will provide both the SRDP equipment—a scanner, a digital printer for pages, a four-color digital printer for covers, and a binding machine—and the staff to produce an annual output of about 36,000 books initially. (If Edwards Brothers' service proves unsatisfactory, we retain the option to bring in another vendor or run the operation ourselves.) The CDDC, in service to participating presses, will manage every other aspect of the use of this production system, from arranging for the creation, storage, and migration of digital files to decisions on timing and size of printing orders to payments to Edwards Brothers to charges to presses and purchasers to shipping orders and storage of small numbers of book units.

Among the advantages to the participating scholarly presses in having the digital book production center and the repository of digital books at the CDC are:

- CDC staff will manage the flow of work to and through the center and thus ensure that its current standards for high quality and fast turn-around are maintained.
- Presses will save both time and money by having the production facility located where their books are stored and shipped.
- The CDDC will be able to experiment with course packs and similar offerings that combine parts of books and journal articles, as scholars request them.
- The CDDC and consortium presses will be able to work with printing vendors in geographically advantageous locations, e.g., on college campuses or in Europe or Australia, as demand warrants in the future.
- If the relationship with Edwards Brothers should cease to be attractive, we can easily turn to another vendor or set up our own digital book production center if the operation is in our own facility where we can learn about its operation from the outset.
- University presses will have a stable, committed vendor of SRDP services in the CDDC. They will no longer be forced to relinquish their books or files to commercial intermediaries whose pricing reflects an early loss-leader strategy and deceptive transfer pricing between their printing and wholesaling operations.

- This complements the activities of the CDC and its client presses in working together to provide books to wholesalers, retailers, libraries, and scholars.
- This enables the participating presses to become thoroughly familiar with this technology and its potential in their production processes.

SRDP PROCESS AT THE CDDC short-run digital printing requires a digital file for a book. This file is created either from the typesetting file for the book, if it is available, or by scanning each page of a book into a 600 dpi bitonal TIFF image. A TIFF file of the book jacket or cover art is used to print the paperback cover. Our plans for creating these book files and maintaining them in the BiblioVault are detailed in the description of the BiblioVault that follows.[15] Initially Edwards Brothers, our SRDP vendor, will create the files for the covers if those files do not already exist. Creating these files is labor intensive and costly ($50 each), so we will investigate alternative methods during the course of the project.

If the TIFF files for books that will be produced by SRDP are prepared in advance and stored, an SRDP order can be completed and shipped to a customer within three days in a closely managed system, as the CDDC will be.

Lain Adkins, associate director of operations, will manage the CDDC. His level of effort will begin at 40 percent in the first six months and decline to 25 percent in the last six months. A full time CDDC production coordinator (to be appointed) will work with Edwards Brothers, the vendor providing the SRDP services for the CDDC, with the consortium presses, and with the Digital Library Production Service at the University of Michigan (our vendor for processing backlist books) to ensure that the CDDC works smoothly. This person will be responsible for selecting, inspecting, and sending books to the DLPS for conversion.

In order to make SRDP an effective strategy for scholarly presses, the CDDC will devise and implement an efficient set of management systems for dealing with the printing system, the client presses, and the warehouse and fulfillment components of the CDC. Some of these management systems will be components of the basic fulfillment system at the CDC; CDC programmers who are expert with this system will build these new components in Fall 2001.

SRDP is used for books with small sales and, hence, modest revenue

15. In the first year as the initial development of the BiblioVault proceeds, we will maintain a set of the SRDP files for the scanned books with Edwards Brothers. Thereafter, these files should be downloaded from the BiblioVault when a title is printed.

potential. Presses cannot afford to spend time making individual production decisions on such titles. Process issues that will be handled in these management systems include:

- Pinpointing backlist titles appropriate in inventory and sales to enter the SRDP program for scanning (a publisher will make the final decision on whether its books will enter the program).
- Creating automated systems for selecting titles to be printed via SRDP when stock is low.
- Devising financial formulas for determining appropriate SRDP print runs and pricing.

This proposal seeks funds to cover the costs of the CDDC manager, the production coordinator, the programming described above, a consulting technical writer to prepare instructions and manuals for cooperating presses, the conversion of books so that they can be processed by SRDP, and office systems and administrative expenses for the production coordinator (funding requested for the BiblioVault is detailed below). The University of Chicago Press will contribute the cost of the time invested by Paula Duffy, Don Collins, and Mary Summerfield, the space required by CDDC staff and the digital book production center, and the cost of building modifications required.

OUTCOMES OF THE CDDC PROJECT Outcomes from the CDDC's SRDP effort will include:

- Systems for managing the flow of books into SRDP.
- Improved finances for the participating presses.
- A better understanding by university presses of the economics of producing and distributing books and appropriate strategies for various types of books.
- Economic analysis of appropriate use of offset printing, SRDP, and POD.
- Economic analysis of appropriate pricing for books based on method of production, risk assumed, and the like.

The analyses will be available not only to the publishers who participate in the CDDC but also to other scholarly presses as we will disseminate them through industry meetings and reports published on a CDDC Web site.

THE BIBLIOVAULT: AN ELECTRONIC REPOSITORY OF BOOKS FROM UNIVERSITY PRESSES

The BiblioVault will be a repository of digital book files and related catalogs for not-for-profit scholarly presses. It will include and deploy as instructed by the consortium presses: (1) scanned page image files for backlist books for which electronic files are not available—a pristine master file, a version to be used for SRDP, and a third version to be viewed online; (2) text files of such backlist books created through OCR processing to be used for online searching and delivery of search results; (3) the existing electronic files for recently and newly published books; (4) metadata about the contents of each book; (5) a technical catalog of data about the book files; and (6) a database about the rights profile for each book.

The BiblioVault will be a strategic intellectual property management tool for these presses as we move into a future in which electronic files are increasingly important assets for publishers. In the future, the BiblioVault will also be a prime outlet for providing scholarly works to the community in an online format, through libraries and directly to scholars whose libraries have not purchased access to its contents. The BiblioVault's most modest online offering will be a search of the contents of its books. The search will result in retrieval of the sentence or paragraph where the requested term is used. This will allow readers to locate books that are relevant to their interests.[16] We believe that this search capacity can be provided well before all electronic rights have been obtained.

Today, scholarly presses struggle with strategic decisions related to the production of their books; they seek maximum flexibility in the use of electronic files in an unknown future. The CDDC will work with member presses to think through their options as they evolve. Our ultimate goal will be to create collections of electronic books that scholars can view and search using multiple tools at reasonable prices while allowing the scholarly publishers to reap a reasonable return on their investment so that they can continue to publish and disseminate scholarly works. Having a substantial repository of digital scholarly books will provide university presses with the ability to work with libraries and scholars to create and offer collections of these books.

This proposal requests funds to support the following elements of the BiblioVault as it is launched over an 18-month period:

16. If a particularly valuable book has rights clearance problems, it would be included in the "search" database only. Alternatively, contents, e.g., images, for which rights cannot be cleared, might be excluded from the elements of the book that BiblioVault users could view.

- The BiblioVault general manager, who will be hired at the end of the first six months of the program.[17] This person will have strengths in both publishing management and information technology. S/he will be responsible for managing the development of the BiblioVault, working with publishers to place materials in the BiblioVault, and working with the marketplace to make optimum use of those materials in the scholarly community.
- The information technology manager, a senior manager at the University of Chicago Press. The IT manager will work on the BiblioVault at a 50 percent rate during the first six months, at a 25 percent rate during the second six months, and at a 20 percent rate during the last six months. Initially s/he will work with Evan Owens and other technical staff, with Mary Summerfield, and with Lain Adkins to design and develop the BiblioVault. When the general manager is hired, s/he will work in concert with the IT manager to continue the development of the BiblioVault.
- A BiblioVault production coordinator who will be responsible for manipulating files for the BiblioVault and for putting information into related databases. This individual will be part of a team and will spend 30 percent of his or her time on BiblioVault activities in the first six months and full time during the last year.
- BiblioVault programming, to be undertaken by UCP programmers or a consultant.
- Rights database consulting and software and prototype rights clearance activity.
- Hardware and related maintenance (to work with current UCP infrastructure).
- Software licenses.
- Office computer systems and miscellaneous administrative expenses for the general manager and production coordinator. Travel expenses for the general manager and IT manager (and for the CDDC manager).
- Conversion of the book files that will be part of the BiblioVault (detailed below).

PRODUCTION PROCESSES FOR ELECTRONIC FILES

In developing the CDDC and the BiblioVault, we will work with both backlist books that lack electronic files and with recently published and about-to-be-published books that have electronic files.

CONVERSION OF BACKLIST BOOKS In this 18-month project, we will scan 2,300 backlist books with no usable electronic file to create a 600 dpi bitonal TIFF file; this master image file will be the basis for derivative files to be used in SRDP and for online viewing. These books will also be

17. UCP will contribute effort by Mary Summerfield, director of development and planning, at a 20 percent rate during the first six months to work with the IT manager to begin development of the BiblioVault.

processed with optical character recognition software, resulting in searchable text files. Rich metadata will be created that "describes" the content of the books. The books included in this set will be (1) appropriate for short-run digital printing in size and will not have images that reproduce poorly with SRDP, and (2) either out of stock or about to go out of stock.

The Digital Library Production Service (DLPS) at the University of Michigan Library will create the master image file, the text file, and the metadata for each of these 2,300 books over a period of 12 months within our 18-month project period. Wendy Lougee, associate director of the University Library (for Digital Library Services) and John Price Wilkin, head of DLPS, will administer this conversion program at the University of Michigan. Thanks in large part to the generosity and vision of The Andrew W. Mellon Foundation over the past several years, DLPS has converted millions of pages of historical books and journals for preservation and access purposes. DLPS will provide a conversion cost of $0.30 cents per page for a minimum of 750,000 pages[18] converted over a 12-month period. This cost includes removal of the binding, bitonal scanning, quality control for each image, optical character recognition, and page level metadata as well as management oversight.

In Fall 2001, DLPS and CDDC management will complete the definition of our relationship and processes. We are confident that using DLPS as the conversion center for the backlist books will allow this venture to achieve lower costs than we could reach otherwise. In addition, we look forward to calling on the deep experience of the DLPS staff as we work to develop the BiblioVault.

PREPARATION OF BOOKS WITH ELECTRONIC FILES

For recent books created from electronic processes, participating publishers will submit PDF files to the BiblioVault; these will become the master files for these books. For some of these books, a PDF file already exists. The UCP production manager has found that the publisher can retrieve it from the printer for about $75. For others, the only electronic file is in a typesetter format and that must be converted to a PDF; the average cost for obtaining such a converted file is about $150. (In our budget we have conservatively estimated that 2,250 of the 2,850 newer books we expect to convert will lack PDF files and will need the more expensive conversion.) From the master PDF files, we will derive files for

18. As we anticipate that these books will average 340 pages, this equates to 2,206 books.

use in SRDP and for use in displaying the books online. The latter files will be manipulated, catalogued, and indexed, as necessary for inclusion in the search system.[19]

ELECTRONIC BOOK RIGHTS Many different types of contracts govern the books published by scholarly presses. In all cases, varying rights are granted to the publisher by principal authors and by other parties who retain copyright to portions of text and/or graphic material. The CDDC can digitize these books and provide search capacity without owning the electronic rights. We believe that fair use will allow us to show bibliographic information, a table of contents, and the paragraph or even the full page where a search term appears; we will need to discuss such use with legal counsel expert in the evolving area of copyright. However, if a BiblioVault offering is to display more than summary data, front matter, or minimal excerpts of text to a user, the publisher will need to obtain electronic rights from the primary authors and from any party who owns rights to text or images in the print version of the book. (SRDP is just another form of printing a book; no particular rights or permissions are needed to use this technology to reprint a book.)

Typically, university presses do not have comprehensive databases about the contracts for their books. Developing a collection of electronic books that scholars can use beyond simple searching will include design of a rights database and input of information that describes the rights situation of each title. We will need to obtain electronic rights when they are not included in the book's contract. As a part of this grant, we seek funds to cover the fees for a consultant to aid in the development of a comprehensive rights database and to acquire appropriate software. We also seek funds to cover 100 hours of prototype electronic rights clearance activity for a sample of titles that will be part of the BiblioVault. This will be a sufficiently large sample to allow us to estimate the time and funds required to clear a large collection of titles and input the information into the database. The experience of building this database and the database itself will prove of enormous value to participating presses. Each press could not afford to invest on its own in the development of a database linked to the electronic dissemi-

19. Our Journals Division processes files in this way routinely. We will also assess the system in use by the University of Pennsylvania Library-Oxford University Press History Digital History Books Project. Our cost estimate of $.05 per page is based on a recent report of the costs experienced by that project. (Cf. Roy Heinz, email message to M. Summerfield.)

nation of books. All of us, however, need to maintain comprehensive records and to pay contracted royalties to our authors and contributors.

We will design the BiblioVault so that it will be possible to show only some parts of a book if we cannot obtain rights to all its components. Or if electronic rights are not available, search results will be limited to variables described above. That information alone will be valuable to scholars in helping to locate titles that are relevant to their work.

BIBLIOVAULT VENTURE MODELING One of the BiblioVault general manager's duties will be assessing the potential for member scholarly presses to offer online book services when the needed infrastructure and content are in place (after this project is completed). The BiblioVault general manager will canvass libraries, scholars, and publishers to determine what is feasible and desired by those groups. S/he will undertake financial analysis to determine what offerings will allow the BiblioVault to become self-sustaining. Even the minimum offering of a searchable database entails costs related to creating and maintaining the database, announcing and providing access, fielding libraries' and scholars' queries, and the like.

If librarians and scholars report that books in particular subjects are most valuable in electronic form, we will suggest to presses that they place books in these subjects in the BiblioVault in the second phase of its development. (In the first phase, backlist books will be selected based on their need for short-run digital printing. We expect participating presses to provide a somewhat comprehensive set of current books for which they have access to electronic files.) An initial expectation is that the BiblioVault online books services might best focus initially on books in the humanities and social science disciplines covered by JSTOR, as many scholars in these fields are now accustomed to using resources online via JSTOR. In addition, books are more important to research in the humanities and social sciences than they are to research in the natural sciences.

BIBLIOVAULT OUTCOMES Outcomes from the BiblioVault part of this project will include:

- The BiblioVault: an electronic repository for scholarly books to serve SRDP and online book uses and any future uses that publishers may find for their electronic book files.
- An initial collection of about 2,300 backlist books that would go out of print without access to SRDP as well as 2,850 recent books that will be positioned for printing via SRDP.

- Models of online book services, including business plans — possibly to be launched in a second phase of BiblioVault development.
- Enhanced cooperation among university presses in preparing and understanding the potential use of electronic files for books published in the past and for books that are in production.
- Processes and tools for tracking and obtaining electronic rights to be shared with other not-for-profit presses.

THE BIBLIOVAULT—A TECHNICAL DESCRIPTION

The BiblioVault is a vendor- and application-neutral repository of book content files and a catalog of book metadata whose initial objective is to serve as a staging area for SRDP production. Vendor neutrality implies that it is the prerogative of the copyright holder to determine where the book content files are to be produced and in what manner. The BiblioVault establishes requirements for the format and structure of the information, or data, it houses but does not impose further restrictions on the production of the data. Vendor neutrality further implies that it is the prerogative of the original publisher of the material to determine when, how, and by whom that material is used for SRDP and for other applications. Application neutrality implies that we will design the BiblioVault to accommodate other applications or uses of the book content beyond SRDP.

In the 18-month period of this grant, we do not envision implementing an electronic book delivery system. However, we will design the BiblioVault with an eye toward applications of this sort by ensuring that we can accommodate a variety of data file formats and metadata elements beyond those required for SRDP. During implementation, we will develop prototypes of electronic book delivery systems to prove that our conception of an application-neutral repository is sound. The prototypes will establish that we have captured enough information to control the construction and delivery of electronic book products. The prototypes will not include access control and sales systems, the other major components of a full-featured electronic book delivery system.

TECHNICAL DESIGN AND IMPLEMENTATION ISSUES

DATA MANAGEMENT The BiblioVault requires data management at two levels: first, the collection and maintenance of data files for books in the BiblioVault's repository; and second, the collection of technical and other metadata about the files in the BiblioVault's catalog.

The repository must provide physically secure storage for data files in two distinct locations (buildings) for a continuously expanding collection of data. The servers comprised by the repository must be placed on an isolated network not directly accessible outside of the BiblioVault. Normal data center operations to insure replication and backup of the data must be conducted. The planning phase for BiblioVault thus includes the design of the server installations that satisfy our redundancy and growth requirements, and the development of protocols for verified data replication between the data centers.

In addition to tracking the details of the contents of the repository, the BiblioVault catalog will record technical and other metadata about its holdings. The catalog will store information about how the data files were created, with what software, using which parameters, and to which version of the file format. As the technology evolves, files in different versions within the same family of file format, e.g. PDF, are likely to have different capabilities, newer files capable of being put to some uses for which older files cannot be repurposed. Similarly, the BiblioVault should be capable of storing and recording appropriate metadata for any file format that can be used to drive the outputs the BiblioVault supports. The ability to report on the state of the BiblioVault holdings along these lines will be an invaluable planning tool for participating presses. Our objective in the design phase of the project will be a database design for the BiblioVault catalog that is abstract enough to accommodate the rapid rate of change in this area.

The design of the BiblioVault catalog must represent relationships among potentially multiple data sets for any title. Under one scenario, for example, one set of files serves as the master set from which other sets are derived for use in specific applications: a file set including high resolution cover scans packaged with a print order and staged for SRDP; or PDF files down-sampled for fast delivery over the Web. Some parts of a master file set may be held in reserve for future applications; for example, the OCR texts associated with the scanned images created for backlist titles although not usable for SRDP will be essential to permit full-text searches online of book content and for use in producing electronic book prototypes.

The BiblioVault must be designed with sufficient generality so that we may expand the number and types of uses of the book contents. We will ensure that new metadata elements can be added without requiring the restructuring or reengineering of the entire catalog.

DATA FILE REQUIREMENTS AND DEPOSIT PROCEDURES

The BiblioVault general manager will work with SRDP vendors to develop format specifications for data files suitable for use in SRDP. The general manager will establish, in coordination with electronic media managers at participating presses, procedures for depositing data into the repository.

It will be necessary to provide documentation of the specifications and procedures for participating presses. We will contract with a technical writer to document this and other aspects of the project.

Because data file formats and BiblioVault applications evolve over time, it will be the responsibility of the BiblioVault's general manager to stay abreast of the current developments, to organize revisions of procedures, specifications, and documentation as appropriate, and to communicate these changes to participating presses.

CATALOGING AND QUALITY ASSURANCE

Whenever data are deposited in the repository, we must verify that we have received a complete set of files, that the files are readable, and that they satisfy the BiblioVault's technical requirements. When placing a file set into the repository, we will verify and catalog the technical specifications of the files. UCP staff with the job title "electronic publishing (epub) specialist" will be responsible for tasks of cataloging, quality assurance, and placing files into the repository; this person is described as the BiblioVault Production Coordinator herein. These tasks are essentially the same as those required to process data files in our electronic publishing operation. By assigning BiblioVault database wrangling tasks to our pool of epub specialists, we assure that we will have staff on hand to do this work when it needs to be done even though we estimate that the work requires a single FTE's effort at the peak of file handling activity during this project.

The BiblioVault general manager will oversee the development of the processes and procedures used by the epub staff and provide for their training. The epub staff will report to the Press's electronic publishing operations manager.

The BiblioVault general manager will act as the liaison with participating presses as general problems arise and as technical specifications evolve.

DELIVERY ORDERING AND DATA STAGING

Although one outlet for BiblioVault material will be the SRDP facility we will establish in the CDC, we believe that the BiblioVault must be designed

to provide SRDP data to other SRDP facilities and that it should be the prerogative of participating presses to decide at which SRDP facility the printing should occur. We will provide a means for staging the delivery of data to external SRDP facilities initially through an FTP (file transfer protocol) site. The BiblioVault will include programming to support delivery requests, authorization, and reporting.

ONLINE REPORTING The BiblioVault will issue a suite of reports for participating presses that includes a listing of titles stored in the repository, their metadata, and their delivery statistics. The BiblioVault general manager will develop this suite of reports in consultation with participating presses.

BOOK CORRECTIONS/EMENDATIONS/ UPGRADES/OBSOLESCENCE It has long been the practice of publishers to include minor corrections to books when new impressions are manufactured. The BiblioVault will accommodate this practice by allowing presses to deposit data in the repository that supercede current data. The catalog will track versions of the data. Similarly, presses may wish to provide upgraded versions of their data files that incorporate the new capabilities of a new version of the format or they may want to provide data files in new formats. Version control is a design goal of the repository and its catalog.

Under some circumstances, data files stored in the BiblioVault may become obsolete. A press's rights to a title may expire. A new edition of a work may be published. The BiblioVault will place the management of the full life cycle of data within the control of the press that owns the data.

INTEGRATION WITH AUTOMATED SYSTEMS

A goal of SRDP and of the BiblioVault is to reduce the overhead costs of initiating a reprint of a book. The process of initiating a reprint can be automated in an order fulfillment system that monitors inventory levels, and that tracks current orders for books and their sales history. As reprints are requested, the BiblioVault must stage the books appropriately for use in an SRDP facility.

Because the CDDC is intended to serve the larger university press community and not only CDC client presses, the BiblioVault technical manager will need to work closely with non-CDC client presses to develop specifications for and to implement the integration of the BiblioVault with other fulfillment systems. The interface defines the information used to control the release of data to SRDP

facilities. The interface may also include manufacturing requirements—what quantity in what timeframe—and shipping specifications—which customer under which purchase order and invoice number via which shipper. The details of the interface will depend on the role played by the BiblioVault in the production process. Major questions arise: Should the BiblioVault serve as an intermediary, conveying not only the data used by the SRDP producers but the authorization to print and other information traditionally represented in a manufacturing purchase order? Should the BiblioVault act only as a server of data? Choices will be based on our ability to serve the needs of the three parties involved: SRDP suppliers; the BiblioVault; and individual university presses. We expect at the end of the grant period to propose solutions based on early experience and extensive consultation with suppliers and presses.

The groundwork we lay for this level of integration should serve us well as SRDP facilities proliferate, enabling remote production of books in kiosks on campuses and in bookstores. While these facilities are presently barely in the pilot stage of development, we will design the BiblioVault with an eye towards participating directly in these processes rather than relinquishing control of our content to commercial enterprises servicing these new sales channels.

CDC FULFILLMENT SYSTEM WORK Integration of SRDP with the CDC fulfillment system is necessary for UCP to reduce overhead costs for itself and for CDC clients. The CDC presently provides UCP and its client presses with inventory management tools based in part on a rich representation of the sales history of the titles it handles. To achieve our goal of reducing overhead costs, we must modify the CDC system to trigger SRDP reprints according to the "business rules" we establish. A typical business rule would specify a quantity to be produced beyond the quantity required to satisfy the orders presently on hand, based on the sales history of the book (otherwise known as the economic reorder quantity). The rule would also assign a SRDP processing priority based on a required ship date, if stated, or on the customer's market segment. Adoption orders for use in the classroom receive a higher priority than orders shipped to wholesalers, for example. The order to SRDP would also indicate which portion of the run is to be staged for packing and shipping and which portion staged for shelving in inventory pick zones.

After a SRDP reprint is triggered and processed, we must capture the quantity and cost of the inventory produced, and then load these values into the CDC system in order to maintain accurate inventory and cost-of-sales records.

Two full-time programmers under the direction of CDC management currently support and enhance the CDC system. They will be assigned to modify the system to automate SRDP processing. We have budgeted $25,000 for this work. The work will take place as the SRDP facility is being installed in the CDC warehouse and will be completed within three months.

TECHNICAL MANAGEMENT PLANNING A senior UCP IT staff member will be appointed BiblioVault technical manager and assigned to work with the CDDC manager and with Evan Owens, UCP IT manager, to develop technical plans for the project. S/he will be charged with overseeing and coordinating the design and implementation phases of the project. S/he will report to the BiblioVault general manager as well as to Evan Owens for his work on the BiblioVault.

The BiblioVault technical manager will oversee the design of both the data centers and the repository and catalog databases in the first two months of the grant period. Make or buy decisions for the repository and catalog software will follow the initial design phase. Implementation will be begin at the start of the second quarter of the project and continue for six to nine months. Within the first three to six months of the implementation phase, we expect to have completed work on the BiblioVault repository and catalog. In the final three months, we will implement electronic book prototypes.

The BiblioVault technical manager will oversee implementation, either by coordinating the programming effort here (if our decision is to "make", or build proprietary software) or by acting as the Press's project manager for the installation of software we purchase. We estimate that a project of this scope in this timeframe can be implemented as part of the work of our team of five programmers and systems analysts, under the direction of the BiblioVault technical manager and the Press's IT manager. As reflected in the proposed budget, the Press is prepared to allocate some of the time of five staff programmers to the project in this period.

As the programming phase nears completion, we will contract with a technical writer to produce documentation for BiblioVault users.

In the second half of the grant period, data prepared under this grant will be loaded into the BiblioVault after processing for quality assurance and cataloging under the general direction of the BiblioVault general manager.

CDDC and BiblioVault—Organization for First 18 Months

STAFF:

Paula Duffy, director, University of Chicago Press, Chair of the CDDC and BiblioVault Management Committee

Don Collins, University of Chicago Press CFO; CDC director, Member of the CDDC and BiblioVault Management Committee

Mary Summerfield, director of business development and planning, University of Chicago Press, and project director—to oversee the overall project and to direct the BiblioVault effort initially.

Lain Adkins, associate director of operations, University of Chicago Press—to manage the CDDC, oversee its staff and vendors, bring presses into the consortium, and manage those relationships. To prepare various reports on the CDDC, economic analyses, and recommendations to scholarly presses, including consortium members.

TO BE HIRED OR ASSIGNED:

BiblioVault technical manager—to work with CDDC management, the BiblioVault General manager, and with other IT staff and consultants to design and develop the BiblioVault. Activities are described at length in the technical description of the BiblioVault.

BiblioVault general manager: This individual will work with CDDC management and information technology staff on the design and development of the BiblioVault as a repository for books. S/he will work with presses so that they understand what steps they must take in order to contribute to and use the BiblioVault, how they need to communicate to their production editing staff, typesetters, and printers, and how to prepare their books files. S/he will work with presses, libraries, and scholars to think about which

online books services might be most worthwhile in a second phase in terms of their utility to the scholarly community and their financial viability. S/he will undertake various analyses of the BiblioVault and propose plans for introducing online books services.

CDDC *production coordinator*—to work with Lain Adkins, with Edwards Brothers, and with participating presses to ensure that the SRDP operation works smoothly and to pull books to be sent to Edwards Brothers for scanning, to gather data on SRDP for analyses, and to keep general records of activity.

BiblioVault production coordinator—to work with CDDC management, especially the BiblioVault general manager and technical manager, on managing the contents of the BiblioVault. Duties are described at length in the technical description of the BiblioVault.

TECHNICAL STAFF

The University of Chicago Press Information Technology Group supports the electronic publishing operations of both the Books and the Journals divisions of the Press. Staff members who will contribute to this project and their areas of expertise are listed below:

Evan Owens, Press IT manager (online journals publishing, SGML/XML architectures, project management)

John Muenning, lead electronic publishing developer (SGML, editorial systems, data conversion, Web design)

Michael Boudreau, senior electronic publishing developer (Web design, DOI, web applications, data conversion, graphics)

Roy Bixler, senior systems programmer (e-commerce applications, system administration, Web site administration)

Sara Zimmerman, electronic publishing developer (search engines, data conversion, Web statistics)

Our operations team of electronic publishing specialists will handle the day-to-day content management activities of the BiblioVault. The team will be increased by one staff position (BiblioVault production coordinator) for this project.

CDDC and BiblioVault—Projected Calendar Of Activities

DIGITAL BOOK PRODUCTION PROCESSES

- Recruit presses to participate in CDDC and BiblioVault—Spring 2001 through Winter 2002
- Develop procedures with University of Michigan Library Digital Library Production Service as provider of conversion services for backlist books — Fall 2001; Implementation to begin by January 2002
- Set key characteristics of scholarly books from university presses to include in BiblioVault initially—Fall 2001

SRDP PRODUCTION

- Continue discussions with Edwards Brothers toward establishing production center at CDC and firming up processes and pricing, including necessary characteristics of files for SRDP (upgradeable over time)—Summer 2001
- Modify CDC for SRDP facility—Summer 2001
- Start-up production equipment at EB Ann Arbor facility for testing—Fall 2001
- Recruit or assign CDDC Production Coordinator—Fall 2001
- Transfer existing University of Chicago Press SRDP files into CDDC repository (also for other presses as they come into CDDC SRDP program)—Fall 2001
- Develop program of services and pricing to cover costs of operation, including setting initial strategy on SRDP runs (minimums, usual rules)—Fall 2001
- Undertake formal discussions with CDC members and other not-for-profit presses about SRDP services — Summer 2001 - Summer 2002
- Develop program for selecting titles from backlist for consideration—Fall 2001
- Develop program for triggering production decision and initiating orders –Fall 2001

- Work with Edwards Brothers to develop electronic data interchange (EDI) for billing and payment—Fall 2001
- Move SRDP system to CDC and start-up with EB staff operating center –Fall 2001
- Scan and create new SRDP files for books as appropriate titles are identified for which supply will be depleted within one year—Winter 2002-Fall 2002

THE BIBLIOVAULT: AN ELECTRONIC REPOSITORY OF BOOKS FROM UNIVERSITY PRESSES

- Assign BiblioVault IT manager—Fall 2001
- Undertake formal discussions with CDC members and other not-for-profit presses about participating in The BiblioVault—Fall 2001
- Undertake technical design and development of the BiblioVault as detailed in the technical description—Fall 2001 - Fall 2002
- Design BiblioVault data centers, repository structure, and catalog database—Fall 2001
- Install BiblioVault data centers—Fall 2001
- Implement BiblioVault catalog—Fall 2001 - Winter 2002
- Recruit BiblioVault General manager—Fall 2001 - Winter 2002
- Develop SRDP specifications for BiblioVault delivery—Winter 2002
- Begin loading data into the repository and populating the BiblioVault catalog—Spring 2002
- Prototype online book search and delivery systems—Summer 2002 - Winter 2003
- Evaluate the interest in providing collections of electronic scholarly books (publishers and authors) and acquiring and using such collections (libraries and scholars)—Summer 2002 - Winter 2003
- Analyze potential demand and costs of providing such collections—Winter 2003
- Prepare final reports on the start-up of the CDDC and BiblioVault for the Foundation, consortium presses, and the scholarly publishing communityWinter 2003