



361 Sinclair-Frontage Road
Milpitas, CA 95035

BACKFILE CONVERSION REPORT PRESENT TO
THE LIBRARY OF CONGRESS

Submitted
January 28, 2005

On behalf of ATI-Imaging Inc., we welcome the opportunity to perform the paper and microfilm conversion projects for the Library of Commerce in conjunction with ISS.

ATI is one of the few companies that can meet the challenge because we have the latest in technology, skilled staff and proven background to provide the optimum results. As an example, we just completed a scanning project consisting of over 280 million images, which was accomplished in under 75 days.

The following information gathered from the onsite visit and analysis thereafter supports this and I look forward to establishing a solid long-term business relationship.

Respectfully submitted:

Tom Windfield
ATI, Inc
800-536-2212 X322
TWindfield@ati-imaging.com

Brochures and test files enclosed:

This document is released by ATI, Inc. for the sole purpose of providing product capability and pricing information to the **CLIENT** and must be considered confidential. All information contained in this document, or obtained in subsequent discussions with ATI is considered confidential, and reproduction or disclosure of product capabilities, pricing, or other material is, unless otherwise noted in this Response, strictly prohibited.
ATI, Inc. ©2005

Table of Contents

PART 1: EXECUTIVE SUMMARY3
PART 2: ONSITE VISIT4
Personnel and Equipment:.....4
Scanning Process and Findings of Library cards:.....4
Findings:5
Indexing analysis of Library Cards:5
PART 3: EVALUATION OF MICROFILM SAMPLES 5
PART 4: RECOMMENDATIONS6
The ATI Approach:6
Scanning and Capture of Library Cards:6
Post Process after Scanning and Capture: 6
Equipment and Staffing:6
Indexing of Library Cards:.....7
Technologies to assist in Automation of Library Cards:7
Methodology in handling the Record Book Images (Microfilm):8
PART 5: ESTIMATED PRICING9
Library Cards:9
Record Books - Microfilm:.....9
PART 6: SUMMARY10
Library Cards:10
Record Books - Microfilm:.....10
Indexing Recommendation:10
Pricing Estimates:10
Flow Chart for the Library Card Project.....11



PART 1: EXECUTIVE SUMMARY

ATI – An Overview

ATI Inc. was incorporated in Delaware in 1979.

ATI, Inc. (ATI) is a privately held global company with operations in the United States, Europe and India, providing a comprehensive range of document solutions designed to meet the needs of document intensive organizations such as government, education, financial services, insurance and legal markets.

ATI formed its Imaging Services Division in 1990 after developing an Imaging System that processed large backfile conversion projects. Some of the largest government conversion contracts in the United States have been awarded to ATI.

We employ approximately 140 people throughout its US operations with additional staff overseas. A very experienced and skilled management group controls each project until completion and also has a proven methodology for each task. A thorough “Statement of Understanding” along with specific training of the production staff takes place for each project. Quality control stages exist throughout the project to gain optimum results. This methodology exists throughout all production environments including our offshore facilities where a significant amount of indexing takes place.

Our equipment inventory available for conversion is significant and constantly improving. The latest technology is utilized and routine scheduled maintenance procedures take place during each shift to ensure desired performance levels. ATI has the capacity to capture over 5 million pages of information per day through its range of technology. Indexing and key verification of scanned documents can be performed on 50 workstations located at the various conversion sites.

In addition to a fully staffed and equipped production facility, ATI specializes in on-site conversion projects. High-speed equipment along with mobile Management and Production teams allow numerous large projects to be done throughout the United States at the customer site to insure the safety and confidentiality of the documents.

Company Mission Statement

ATI is committed to providing the highest quality products and services for the lowest possible price, without compromising the value of the final product. With a commitment to personalized service, specialized project management, highest quality equipment, cost-efficiency and each client’s overall satisfaction, ATI has become a leader in the field of document imaging. Finally, maintaining a focus on our “Mission Statement” has allowed ATI to grow significantly in size and capacity that will allow us to accomplish the project requirements of the Library of Congress.



PART 2: ONSITE VISIT

Personnel and Equipment:

Tim Plette, VP of Production, James Kerkstra, Project Manager and Tom Windfield, Regional Manager arrived on site bringing with them the following equipment:

- 1- High Speed Canon 9080 C - Color scanner
- 1- Dell PC and Monitor with appropriate scanning software

The team arrived on November 30th and set-up, obtained the reader cards and completed the entire required scanning process of all 27 drawers by end of day December 1. In addition, we were escorted throughout the file rooms and viewed the microfilm as well as bound material that will eventually be part of the project. On Wednesday morning, Tom Windfield met with Baha Akpinar to gain further insight and direction of the overall LOC conversion project.

Scanning Process and Findings of Library cards:

The following summarizes condition of each card drawer:

<u>DRAWER</u>	<u>CONDITIONS-FINDINGS AND NOTATIONS</u>
Applications	Pink - all cards required unfolding and refolding (folded 6.5x6) Double sided
Prints/Labels	Some tissue cards single sided (3x5)
1946-54	Cards thicker than normal, single sided (3x5)
1955-70	Same card type and thickness found previously, single sided
1971-77	Same card as in previous 2 drawers with some staples - single sided
Assignments	Same thickness and size – some yellow, pink, and tan – single sided
Combined	Same thickness & size – multicolor throughout (blue, yellow, pink) single sided
Notice of use	Same thickness and size – off white – single sided
Notice-intent	Same thickness and size – blue – single sided
Bks (Author)	Same size – off white. Difficult to feed (info taped on card, some double sided)
Bks (Claimant)	Most 6.5x6, some smaller, double sided (unfolding/refolding) – difficult scan
Graphic Arts	Some variation in thickness (slow feeding) size of 3x5 – off white – single sided
Period. (C&T)	Approx. 3x6 – off white with some green cards, slow to feed, single sided
Period. Title	3x5 off white, light text, difficult scan (thickness) both single and double sided
Period. Title	Same images and condition of cards as drawer 13
Dram. (Claimant)	Approx. 3x6 – multicolor, slow feeding through scanner, double sided
Dram. (A&T)	4x6, multicolor with some carbon – single sided
Gen. Index	3x5 both colors and off white – some thin cards, difficult feeding – single sided
Music	3x5 in size, standard thickness – double sided
Pseudonyms	Smaller size, some double sided, some with blank fronts and info on the back
Renewals (3)*	3x5 (Approx 2 out of 50 required unfolding and refolding)
Periodicals	3x5 - Pink cards and some were blue carbons
Graphic Arts	3x5 - Pink cards and some were blue carbons (tested bitonal and gray scale)

* Summary of all 3 Renewal drawers, which had very similar cards in size and color



Findings:

During the scanning process, we experimented with color, bitonal and gray scale settings throughout the entire procedure to gain the best result due to the fact that the cards varied so much in color, size and print quality. Also, brightness settings were constantly being changed in an attempt to capture light colors. Example of setting changes were:

DRAWERS

LOC 10 & 12

LOC 13 & 14

LOC 16

ADJUSTMENTS

Brightness to 220

Light brightness to 90 - 256 gray scale

Because of carbons, used brightness of 190

Indexing analysis of Library Cards:

Two experienced people from our staff, who also have technical background, have been indexing the scanned Library Cards from the on-site visit. Because of the different periods in time and origination, there is a significant lack of field consistency. The positioning of data on the cards are commonly located in different areas. Often times there are multiple entries associated with a field and the data can exist in different languages. This data is found in both typewritten and handwritten entries. All of these aspects often cause the indexer to read and interrupt the card before being able to key the fields in correctly. During the indexing of the scanned cards, many formulations took place to establish the best practices both for manual keying assisted by automated methods.

PART 3: EVALUATION OF MICROFILM SAMPLES

The test consisted of 16 rolls of 35mm and 13 rolls of 16mm microfilm. The quality of images ranged from acceptable to poor. Acceptable images could be read without magnification and the very poor ones could not be read at all. The viewing variables related to such aspects as size and nature, registration, and type of print along with background and speckles. In addition, simple indexing fields were consistent and recognizable. The majority of film tested was acceptable.

The following is a representative evaluation:

<u>REEL</u>	<u>DESCRIPTION</u>	<u>CONDITION</u>
27-161	Record Books (Entries of Title) Yr. 1907	Large hand written documents, average quality requires zoom-in for most, decent Reg. Number
111-104	Application for Registration Of Claims Yr. 1973	Most difficult to view, legibility with manipulation, Common indexing fields - Class and Reg. Number
195-004	Application for Registration Of Claims Yrs. 1950-1962	Better to view, some background and bleed, Common indexing fields - Class and Reg. Number
155-085	Assignment Title Cards Yrs. 1927-1956	Good legibility, some background and speckles Common indexing fields



PART 4: RECOMMENDATIONS

The ATI Approach:

With many years of experience, a methodology has evolved and been developed for dealing with large backfile conversion projects. ATI will automate as much as possible, practically and economically. Best practices have been developed for these factors:

- Separation of work done on and off site for efficiency
- High Speed Internet Access if available and secured
- Indexing locations both at our facilities in the US and offshore
- Quality of documents and the number of form types
- Indexing – number of appropriate fields and length for desired search ability
- Aids to help automate the process from related databases (computer assisted retrieval)
- Use of different Technologies

Scanning and Capture of Library Cards:

Scanner Type – During the onsite test, we used the Canon 9080 C Color Scanner. It was found to be durable and possess the necessary characteristics to capture most source materials with the exception of the smaller size cards. Selecting the correct scanner is imperative to ensure a successful conversion. Due to the varied thickness and different sizes of the cards, we will use a larger and full-featured scanner designed to handle extreme applications. It is the Kodak i600 Color Scanner series that has enhanced speeds, expanded flexibility in quality adjustments and maintains a very durable feeding system that can automatically adjust to different stocks. It also has easy operator access to the feed transport area for enhanced efficiency.

Capture Process – Because of the volume of cards, it is imperative that optimum quality is gained the first time. Rescanning of cards would be very cumbersome and costly. ATI plans to scan in 24-bit color using the JPEG 2000 standard compression technology. The selected Kodak scanner allows color scanning to be done at maximum speed to maintain efficiency. Scanning cards in color allows the images to be manipulated in bitonal or gray scale. Also, other practices such as full OCR can take place for maximum results. The JPEC 2000 technology is extremely efficient and allows presetting the compressed image size. A library card can be scanned in 24-bit color and compressed to a 50 kilobyte or less size to achieve a photograph-quality image.

Post Process after Scanning and Capture:

After scanning the images in color, proper rotation, cropping, deskewing, blank page elimination and quality enhancement techniques will take place attaining the most optimum image to index.

Equipment and Staffing:

The number of scanners to be used for this project will be based on the hours of operation and scheduled deadline. Scanners are rated at 80 to 120 pages per minute. In an optimum environment, this equates to 4,800 to 7,200 scans per hour. Likely, no more than 4000 scans per hour can be expected.



Equipment and Staffing – Continued:

To achieve this rate, we anticipate workstations to consist of a scanner and 4 key staff members to include:

- 1 Prep person
- 1 Scanner operator
- 1 Post-prep person
- 1 Quality Control and overall workstation Supervisor

Our methodology is to provide a prep and post-prep person in order to keep a consistent flow of cards to the scanner, which will increase the image output. We are accustomed to providing multiple workstations in order to complete the project within the predetermined timetable.

Indexing of Library Cards:

We believe that focusing on the primary fields is the best practice. The 7 key fields include: Registration Number, First Registration Date, Title, First Personal Author, First Corporate Author, First Personal Claimant, First Corporate Claimant. An additional open field that you refer to, as “Data Notes” will be tested for automatic population using Forms Recognition and OCR technology. A quality result is somewhat of a concern in this area.

Because of the variations of data, it is our determination that the 7 primary fields, done manually will encompass approximately 100 characters per card with an input rate of approximately 2 cards per minute by a trained operator. Based on this our recommendations are:

1. Offshore indexing – ATI can assemble a large staff of skilled and supervised people to key fields from our offshore production group maintaining a high level of security for approximately 50 percent less cost compared to doing it in the US. An organized and well-managed process will be set in place with the very best Q and A practices. We currently do this for other major organizations with highly sensitive data.
2. One area to assist in auto-population of data is to have a Prep person at time of capture record the label of each file drawer.
3. Enter the first item listed on a card only pertaining to Author, Date or other entry.
4. Utilize correct technologies to automate – (Forms Recognition with limited OCR)

Technologies to assist in Automation of Library Cards:

ATI utilizes every current and meaningful technology in the backfile conversion projects that can enhance or automate the process. The document/forms application can be modified after the scanning and capture process to optimize the result with successful use of secure and highly accurate document recognition methods including Bar Coding, Optical Character Recognition and Information Character Recognition. With the nature of the Library Cards, limited results using OCR is expected. However, tests will be done to explore all accuracy and economics.

Some benefits can be obtained from ‘Forms Recognition’ where we can identify the form and form-layout allowing us to extract and present the required indexing fields to the indexer in the same position on the screen for keying. For example, an Author’s name may appear in a different position on every library card but we intend to present it on the indexers screen in the same position every time to assist in the keying process – this will help minimize the visual search. By recognizing and extracting the fields in this manner, we will achieve greater accuracy and more rapid indexing.



Technologies - Continued

Other benefits from using Forms Recognition may be in the area of securing a successful OCR read of some fields. In recognizing the form, it is possible that we can automatically extract and repair the characters and attempt an OCR read. If the OCR has the desired result and has properly recognized the field, the indexer can view and comparison with the original image. However, our experience has been that if there are 3 or more failed characters due to a poor OCR result, it is best for the indexer to key the entire field. Upon review of all data, we believe that a Forms Recognition process will increase speed and accuracy in the keying of indexes. However we are guarded in our expectations that OCR as part of the Forms Recognition process will provide much benefit.

Methodology in handling the Record Book Images (Microfilm):

In conversion of the Record Book Images, there are 3 options:

- Direct scanning of Books –on site with an overhead scanner
- Unbind the books – on site with a production sheet scanner
- Scan the microfilm

Of these options, film scanning is preferable and will be much less costly. Because of this we believe that as much film and likely all should be scanned and indexed especially if the objective is to provide data that leads to access points. Even with the poorest of film, the final results could accomplish the LOC objectives.

In reviewing the quality of the film samples provided, it is estimated that approximately 85% to 90% will produce proper results without much manipulation. The biggest quality issues in the samples appeared to be with the “Applications” which contained small and distorted print. Using the latest in clean-up technology, it is expected that many of the poor quality documents can be recovered to have suitable legibility. It is also noteworthy that each of the poor documents had portions that were very readable. Furthermore, it would be interesting to compare the poorest film with the source images. Possibly those documents are of poor quality and the rescanning may not necessarily provide the desired results.

Scanning - ATI will use the best high quality film scanner available which is the Wicks & Wilson RS 300. We have 15 of these models at our US production facility. It has tremendous automated image and feeding adjustments that will certainly enhance and clean up the quality. This scanner along with very skilled operators will impact the overall quality of the film. Images will be scanned in gray scale to realize a better image using JPEG 2000 standard compression technology.

Indexing – It appears that there will be some standardization with the different types of records. Classification, registration and year seem to be common. Different fields may be required for different types of records but some standardization is possible. We recommend avoiding an entire date, as that can be manually intensive and costly. Technology, previously described, will be used to help with the manual keying process. This may include several automated methods such as ‘Forms Recognition’.

Although film conversion is likely to be the best overall option, ATI can provide the rescanning of bound books in a partial or complete manner on-site if desired.



PART 5: ESTIMATED PRICING

The following estimates are based upon our analysis thus far. This is not intended to be a best and final but rather to provide a solid idea for planning purposes.

Library Cards:

Based on approximately 47 million cards with automated technology used as applicable for enhancements and indexing: (Note – A blank side has no charge)

On site Scanning, Prep and QA described: \$0.055 per image

Indexing based on populating 7 Primary fields and OCR for ‘data notes’ where possible:

(Note – Fewer fields will lower this estimate)

Work done Offshore: \$0.15 to \$0.20 per card

Work done in the US: \$0.30 to \$0.35 per card

Record Books - Microfilm:

There is an economy of scale in film scanning meaning that the pricing stated below is done without true knowledge of the number of images. The following is based on review of the samples and the assumption of indexing fields (such as classification number and year) with automated technology used as applicable for enhancements:

Scanning and Preparation (our facility):

16mm film: \$0.025 per image

35mm film: \$0.045 per image

Indexing \$0.015 per field

(A field is defined as - up to an 8-digit number. If significant alpha keying is required, a test can be made to validate an amount per field)

Other Aspects:

- All travel and lodging is included
- Freight is not included and nor part of this estimate
- Final storage media and associated costs are not part of this estimate
- Changes are welcome and could decrease/increase this estimate
- Any specific terms and conditions are not reflected in this estimate

All estimates are intended to provide a benchmark and can be adapted based on changes keeping in mind that ‘quality results’ should not be compromised! The biggest expense to the Library Of Congress would be if an entire rescan and index occurred.

In going forward, we welcome the opportunity to meet with your team in person to discuss our methodology leading to the pricing estimates.



PART 6: SUMMARY

Library Cards:

ATI is well staffed and equipped to perform the conversion projects for the Library Cards. A skilled-well managed staff is prepared to do the following:

- Prepare a complete and all-inclusive plan for the entire project
- Institute our Production Control Database software that tracks entire activity
- Exercise full Quality Control, Exceptions and Reporting processes
- Perform on-site preparation and scanning of all images
- Reassembly as desired
- Post scanning process to ensure optimum results as well as handling exceptions
- Thorough review of project and proper finalization before leaving facility
- Indexing of predefined fields matched to the database – manually with automation
- Final Quality Assurance review
- Delivery of the indexed digital media and full color images

Please see the attached flow chart (Applicable to film conversion with minor deviations)

Record Books - Microfilm:

Although able to accommodate the scanning of books onsite, we believe that the best value is to scan the microfilm. ATI is one of the best companies in the world in film conversion having the latest in technology along with a trained and qualified staff. The process is as follows:

- Prepare a complete and all-inclusive plan for the entire project
- Institute our Production Control Database software that tracks entire activity
- Preview and Preparation
- Exercise full Quality Control, Exceptions and Reporting processes
- Scan the film using a variety of image enhancements and clean-up technology
- Quality Assurance review including managing exceptions prior to the indexing stage
- Indexing of predefined fields matched to the database – manually with automation
- Final Quality Assurance review
- Delivery of digitalized and indexed images

Indexing Recommendation:

Due to the fact that images are for public viewing, we see no security issues in doing the indexing offshore. ATI manages and controls millions of images being indexed each year and has a well-defined process in the handling, security and quality of work leading to the desired results. Therefore, we recommend using our offshore staff for indexing.

Pricing Estimates:

ATI partners and teams well on large projects of this nature and have in place all the necessary aspects to arrive at the best and final results. ***When selected, we will implement methodology and testing applications to insure that the best overall price possible is derived without sacrifice in quality.***



Flow Chart for the Library Card Project

