

“For a microfilm scanning project of this magnitude, a vendor needs to have the resources to creatively adapt to challenges that inevitably come up. BMI delivered.”

Tim Mitchell  
San Jose State University, ScholarWorks



## INDUSTRY

- Education, Colleges & Universities

## LOCATION

- San Jose, California

## CHALLENGES

- 80 years of *The Spartan Daily* was trapped on physical bound books and microfilm
- Retrieval for patrons was cumbersome, involving appointments to a “special collections” room.
- Lack of indexing made searching difficult for patrons

## BMI PRODUCTS & SERVICES

- Newspaper microfilm scanning
- High resolution, searchable newspaper PDFs

## BENEFITS

- 75% less expensive than scanning physical bound newspaper books
- Online web search of any historical newspaper now available via text-searchable PDFs
- Fast, immediate retrieval using keywords to find articles
- Newspaper browsing maintained with full issues available as PDFs



## Overview

From its beginnings as a school focused on training teachers, San Jose State University (SJSU) has matured into a metropolitan university offering more than 145 areas of study with an additional 108 concentrations.

SJSU's ScholarWorks promotes discovery, research, cross-disciplinary collaboration and instruction by collecting, preserving and providing access to scholarly work created at the university. One archive ScholarWorks offers for free, public access is *The Spartan Daily* — a newspaper produced by journalism students since 1934.

More than 80 years of *The Spartan Daily* was not digitally accessible and was only available as aging newsprint in physical bound books and on tattered and torn microfilm. Researching this history was a challenge. SJSU's School of Journalism and Mass Communications, in partnership with the Dr. Martin Luther King, Jr. Library's ScholarWorks, planned to digitize the physical books and contacted BMI to provide the scanning services. However, BMI presented the alternative of scanning the microfilm archive using their first-generation film positives since they photographed much of the original microfilm. This reduced the project price by 75% and eliminated the need to remove the physical archive from the university and cut apart the bound volumes and page scan the hardbound book archive.

Today, all of the legacy *Spartan Daily* newspapers are easily accessible on the Internet as searchable PDF files. Users can browse legacy issues and search using any keyword (e.g. last name, date).

## Historical Bound Books are Cumbersome for Researchers

*The Spartan Daily* was accessible to patrons via physical bound books for more than 80 years. Tim Mitchell, SJSU, states, “To put this in perspective, imagine that you know your Great Grandfather played football for SJSU in 1942 and you wanted to find out more. Accessing *The Spartan Daily* during this time period would require an appointment to gain access to the Dr. Martin Luther King Jr. Library’s “Special Collections” room, where you’d wear white gloves to search for articles in the bound book(s). It was a tedious and difficult process and the inability to search for keywords like a last name made it challenging to find articles and photos of your Great Grandfather even if they were there.”

## Overcoming Challenges in a Digital Scanning Project

When SJSU’s ScholarWorks originally contacted BMI, the organization was looking to scan the few hundred bound books of original newspapers. However, SJSU ran into some hurdles. First, it was very difficult to understand how much the project would cost because it was dependent on the number of pages to scan. “With any document scanning project like this, it is tough to know how many pages need to be scanned. In the end, our rough estimates would have been 20,000 pages off the actual amount,” Mitchell says.

Secondly, scanning the historical books would have required the bound volumes to be cut apart. There was a risk that valuable information could have been lost along edges if accidentally trimmed. In the end the integrity of the historical book archive would have been destroyed by the scanning process.

## A Better Way: Scan the Microfilm Archive

“We had not thought about the microfilm copy until BMI mentioned this as an alternative approach and quality samples were provided to show how the pages would look,” adds Mitchell.

By scanning the microfilm, SJSU was able to preserve the integrity of the historical books.

The price was also 75% less than the price to scan the bound books. Mitchell continues,

“Proposing a less-expensive solution was a real nod to BMI’s customer-centric approach, and helped us get this newspaper scanning project moving in the right direction.”

By working with an established document scanning company like BMI, the SJSU team was able to work through additional challenges that came up during the microfilm scanning phase.

“We discovered that some of our student publications had printed the wrong page numbers and dates on some of the pages,” recalls Mitchell. “A few newspaper issues had printed watermarks that presented scanning challenges.”

Each time anomalies like these came up BMI was able to leverage software, experience and processes to achieve a clean set of high resolution PDF files for SJSU.

Mitchell states, “For a microfilm scanning project of this magnitude, a skilled company needs to have the resources to creatively adapt to challenges that inevitably come up. BMI delivered.”

## The Completed Solution: Online Text Searchable Articles

Today, *Spartan Daily* newspapers are easily accessible on the Internet as viewable, searchable and downloadable page-image PDFs. Anyone looking for *Spartan Daily* material online may find information on the archives. People all over the world are accessing these records through keyword search engines like Google to go directly to the document pages. Genealogists building family trees, obituary writers, former students and others are all regularly accessing the historical records that were previously difficult to access.

Mitchell concludes, “It’s exciting to see history like this revisited. Nobody knew it really existed when it was trapped on physical newspapers. Now it’s unlocked for anyone to access over the web.”