Title: The Campus Newspaper—From Newspaper to Digital Archive, the Libraries’ Role and Responsibilities

Abstract: The Daily Collegian is an entirely student-run newspaper of The Pennsylvania State University. It has been continuously published since 1851 and has recorded the student life, academic changes, and social and cultural changes as a small “Farmer’s College” became a major American research University. “The Collegian,” as it is called, is an award winning newspaper, which has served as the training ground for future Pulitzer Prize winning journalists. The University Libraries has always preserved the newspaper on microfilm and approached The Collegian staff about digitizing the full-page image, due to strong interest by library users, researchers, and alumni for historical content and access to Collegian issues. It is an interesting story--how step-by-step and with each negotiation--the University Libraries will now effectively serve as the repository for all future issues of this important campus newspaper which has strong alumni interest, particularly as The Collegian, moves to a more digital presence and away from its print format. This paper will present an overview of what issues a library must consider as it seeks to provide a digital archive of a campus newspaper.

Name of Presenter: Debora Cheney, Head, Library Services to World Campus/PSU Online and Larry and Ellen Foster Communications Librarian

Employer: The Pennsylvania State University, University Park, PA 16802

Contact information: 309B Paterno Library, The Pennsylvania State University Libraries, University Park, PA 16801. E-mail: dcheney@psu.edu

Biographical statement: Debora Cheney serves as the Foster Communications Librarian supporting students and researchers in the College of Communications, Penn State University. She has also worked closely with supporting access to news content, published and presented on issues related to newspaper database use by faculty and students.

Language of presentation: English
The Campus Newspaper—From Newspaper to Digital Archive, the Libraries’ Role and Responsibilities

In early spring 2003, librarians from the Pennsylvania State University Libraries and staff from The Daily Collegian met for the first time to discuss a possible partnership that would make the historical issues of The Daily Collegian more widely accessible. The Daily Collegian, like all newspapers, was changing. It had launched a website, The Collegian Online in 1996, but was struggling to convert its historical content to make it accessible to readers and researchers through this website. While the University Libraries also had an interest in making the historical issues of The Daily Collegian more widely accessible, it was uncertain of the status of the post-1922 content and recognized that those issues presented the most interesting content to students and alumni. From a technology perspective, the Libraries knew the goal was achievable, IF the University Libraries and The Collegian staff could agree on whether the project should move forward.

This meeting began an interesting story of how--step-by-step and with each negotiation--the University Libraries now serves as the repository to the Historical Digital Collegian database and provides online access to all issues of The Collegian through 2003. This paper will present an overview of what issues a library must consider as it seeks to provide a digital archive of a campus newspaper.

A History of The Daily Collegian

“Collegian has been the target of protests, picketing, and the ceremonious burning of issues of the newspaper on the steps of Carnegie Building (the former location of the Collegian offices). Staff members have been dismissed and editors and reporters have been suspended. Through it all, Collegian has learned valuable lessons. Sometimes it’s easy to forget that students are behind Collegian, writing editorials which sometimes struggle to understand the world around us.” (Collegian, Inc., 2013)

The Daily Collegian is an award winning campus and student run newspaper which has been the training ground for Pulitzer Prize winning journalists and photojournalists. The “commitment of time and the talent of thousands of staffers on both the news and business staffs” is a “labor of love”. (Krasnansky, 2006a, p. 15)

However, two traditions are embedded in history of The Collegian. From its beginning, the campus newspaper “adopted a liberal editorial policy and operated without censorship from the College administration,” (Caggeso, 2006, p. 25) and, thus, “established a tradition of an independent student newspaper that lives on more than a century.” For these reasons, The Collegian presents a unique and valuable perspective on campus life and university history—it is not a mouth piece of the University and it reflects the opinions and reporting “through eyes that are forever young.” (Krasnansky, 2006a)

Penn State history begins in 1855—“chartered as one of the nation's first colleges of agricultural science, with a goal to apply scientific principles to farming.” (The Pennsylvania State University, 2013) The first campus newspaper, The Free Lance began in 1887--an irregularly published “newspaper-literary magazine.” (Caggeso, 2006, p. 18) The first issue of the Free Lance, which preceded the State College newspaper, The Times, by 11 years (Krasnansky, 2006a, p. 15) was delivered to a student waiting at the train station (12 miles away) and “contained 12 pages and sold for 15 cents.” (Caggeso, p. 18) When The Free Lance “succumbed [17 years later with the April, 1904 issue], it was largely due to the ‘lost interest in the mixture of ‘stale’ news and homegrown literature and poetry. As one editor noted, ‘The March issue came out on April 21 and contained news of February 2.’” Penn

1 Collegian, Inc. was chartered in 1940 “as a nonprofit corporation so that the College could distance itself from the wrath of alumni of those in Harrisburg who control the purse strings and are offended on occasions by the paper’s contents or the views of its editors.” (Krasnansky, 2006a, p. 14)

2 The Collegian is overseen by 3-4 full-time non-student staff who oversees the day-to-day operations and budget of the newspaper. A student managing editors and a business editor are selected from applicants each year. These managers then select their news and business staff from student applicants. In addition, here is an active Alumni Interest Group (AIG) of Collegian alumni.
State had become too big for *The Free Lance* to keep pace.” (Caggeso, 2006, p. 25) Immediately in September, 1904, a “pert weekly newspaper, *The State Collegian*” (p. 25) replaced *The Free Lance*, and has been published—under various name variations and banners and with varying frequency as: *The Penn State Collegian* and *The Daily Collegian*.

> “From its earliest days, the *Free Lance/Collegian* focused on campus life, but found room to comment on events beyond its horizons. As communications improved, the paper devoted more space to state, national and world events, with special attention to news that affected students and Penn State.” (Krasnansky, 2006a, p. 15)

*The Collegian* has recorded the student life, academic, social and cultural changes as a small “Farmer’s College” located in isolated central Pennsylvania became a major research university. (The Pennsylvania State University, 2013a) *The Collegian* offers important perspective and views of changes in American education, the Pennsylvania State University campus, the community it resides near/in; and the impact state, national and international events had on student’s lives. It allows us to see our earliest interests in sports—that specifically, American football (our first team was dressed in pink and black, not blue and white) (Edwards, 1978); reflects how a University and its students responded to social and cultural issues such as women’s rights and diversity; documents the impact of wars on students and campus activity; records town and gown relations; and reports on our ongoing struggles with student social behavior.

**The Daily Collegian: History**

In 1996 *The Collegian* launched its first website and it also raised expectations that, not only would *The Collegian* be able to “deliver news at it was happening.” (Collegian, Inc., 1996) it would also be able to improve access to the content of the historical print issues of *The Collegian* newspaper (through its Search Archives and Browse Issues features) which have continued to be published. Researchers, alumni, and students had only the most limited access to:

- Bound historical issues located in either the University Archives’ Reading Room (located in the University Libraries) or *The Daily Collegian* news room (located, most recently, in a building in downtown State College);
- Microfilm of historical issues located in the Libraries’ Microfilm Department; and
- Indexing and discovery provided by a hand-written index card file (begun in the 1930’s) located in the University Archives Reading Room, and, after 1980, a highly–selective print (and later, an online database) compiled by the Microfilm Department.

Those who wanted to conduct research in *The Daily Collegian*, needed to be prepared to spend some time doing so. Despite the limited access to the microfilm and bound volumes and limited indexing the number of requests to use each indicated that many students and researchers still found *The Collegian* a useful and valuable resource and, as with all newspaper content, were surprised that the historical content was not easily accessible via the Internet. In 2002 when the University Libraries received state resource center funding to create online resources for Pennsylvania citizens, it earmarked a project category for a newspaper project, which would increase access to historical Pennsylvania newspapers and would be made freely accessible to Pennsylvania residents.

Notes from early meetings indicate the focus was on newspapers from three different communities: Harrisburg (the state capital), State College (home town to The Pennsylvania State University); and Penn State University.

---

3 Another exciting chapter in Collegian history unfolded during the summer of 1996 with the launch of *The Daily Collegian Online*. During our first year of online publishing, *The Daily Collegian Online* set new standards for delivering information to the Penn State community. For the first time, Collegian could deliver breaking news as it was happening. (Collegian, Inc., 2013)
Newspapers from each of these communities would present an important perspective on the historical growth and development of the Commonwealth from varying viewpoints. However, several factors were considered:

- the history and publication pattern and of the newspaper(s) from each community;
- copyright and the current ownership of the newspaper(s);
- usage patterns and research/teaching support;
- post-1922 issues available for digitization; and
- whether there were already digitized archives and/or content available from commercial vendors or other sources.

For example, many newspapers have been published in (or near) Harrisburg as early as 1799, but some of these have been digitized by Readex through 1922 or are minor newspapers published for brief periods. The present-day Harrisburg newspaper—*The Patriot-News*—is privately owned and its owners have signaled little interest or willingness to make their newspaper freely-accessible either in aggregator databases or through a freely available digital archive. The pre-1922 State College newspaper, the *Times* (begun in 1898), was not yet digitized by a commercial vendor, but the present day issues are owned by Knight-Ridder/McClatchy and the Libraries provided access to this content from 1995 to present through an aggregator database. Only *the Free Lance/Collegian* represented a newspaper with a long and continuous publishing history, for which content, at least through the creation of *The Collegian Online* website, was largely unavailable online and was owned by a non-profit who might be more open to a digitization project.

**Beginning Discussions and Negotiations**

“The University Libraries is proposing to begin a digitization project of the historical issues of the *Daily Collegian*. As a state resource library, we can provide funding and staff for this project, which we believe will provide a model for future newspaper digitization projects. We are aware of significant interest in the historical content and coverage of this newspaper. —from a letter to Collegian, Inc. general manager and chair, Board of Directors, 2003.

A group of cautious but enthusiastic librarians—the Head, Libraries Special Collections/University Archives, the Head, Preservation and Digitization Department, and the Head, News and Microforms Library—met with Collegian, Inc. full-time staff who oversee the students and day-to-day operations of the newspaper. The librarians pitched the idea of an initial project that would digitize all issues of the *Free Lance and Collegian* through 1922 at the least, or, hopefully, through the first major title change, the *Penn State Collegian*, a semi-weekly and weekly that flourished September 28, 1911 to August 2, 1940. Our reasoning for proposing these years were the pre-1922 issues could be digitized without permission of Collegian, Inc. We still owned paper copies but not microfilm of *The Free Lance* so the state funding of this project would be an opportunity to convert our last remaining issues to microfilm; and the pre-World War II issues were likely of less historical interest to the present day Collegian, Inc. staff as fewer alumni or former Collegian staffers would be alive. It felt it might be a way to get our foot in the door and demonstrate our commitment to providing access to the campus newspaper.

We were aware that if we wanted to set the stage for future digitization projects it would be more appropriate to establish a working relationship with Collegian, Inc. A small beginning project would also help us and them understand how quickly and the complexities that would arise from such a project; allow us to develop workflow and production methods; develop more in-depth knowledge of the Olive software, its segmentation process and accuracy; and develop a database for us to evaluate searching and discovery functionality of the Olive interface.

Initially, the Collegian, Inc. General Manager was highly skeptical that the Libraries could digitize 16,000 page images in a year or less—the projected time table for the project. We were granted permission to move forward on all issues April 1887-August 2, 1940, but we also learned:
The intent of The Collegian Online website was to serve as both an archive and to provide electronic access to the daily news, however:
- the website only included a PDF of the front page of each print issue—no digital equivalent of the complete print edition was available or planned;
- the website included only those news/opinions/editorials/photography by Collegian student journalists, it would not include any licensed or syndicated wire stories/photographs/cartoons, columns or a record of any advertisements or supplementary materials included with the print edition;
- Collegian, Inc. was trying to recover and create a searchable historical archive of content prior to the launch of the website:
  - They were unable to recover any content from 1960-1982 from their Stauffer Gold Library system and were expecting they would be able to use their bound volumes and OCR technology to re-create this content;
  - They were working on converting the student journalism articles-only from the 1988 issues by re-typing this content and had some content already finished;

In return for allowing us to digitize issues 1887-1940, the University Libraries agreed to:
- provide all funding for conversion, software/hardware and on-going maintenance and upgrades;
- create links to The Collegian Online website to the Historical Digital Collegian archive and vice-versa.
- worked closely with The Daily Collegian and Libraries publicity and marketing department to develop promotional materials and full-page ads about how to search the Historical Digital Collegian database (Collegian, Inc., 2004)

This first project was successful in demonstrating how effective the Olive digitization/scanning process was and how quickly the Libraries could manage the production process and make resulting databases available to researchers. The 1887-1940 segment was completed, made publicly accessible, and the first publicity about the project and the database was sent out in spring 2004—about one year from that first meeting with the Collegian, Inc. General Manger/staff. Eventually, Collegian, Inc. granted permission to digitize all issues through 1988 which was completed over four years (see Table 1).

<table>
<thead>
<tr>
<th>Phase</th>
<th>Years</th>
<th>Pages</th>
<th>Availability</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase I</td>
<td>1887-1940 (FY 03/04)</td>
<td>16,332</td>
<td>Available 6/24/2004</td>
<td></td>
</tr>
<tr>
<td>Phase II</td>
<td>1941-1955 (FY 04/05)</td>
<td>48,960</td>
<td>Available 6/28/2005</td>
<td></td>
</tr>
<tr>
<td>Phase III</td>
<td>1956-1976 (FY 05/06)</td>
<td>29,916</td>
<td>Available 7/14/2006</td>
<td></td>
</tr>
<tr>
<td>Phase IV</td>
<td>1976-1987 (FY 06/07)</td>
<td>37,428</td>
<td>Available 1/31/2008</td>
<td></td>
</tr>
</tbody>
</table>

The Total Pages, 132,736  
Total Cost: $178,541.00  
Average cost per page $1.35

Some additional developments and milestones:
- Located 3 missing issues from our microfilm holdings from alumni by posting request to locate on our Historical Digital Collegian Website
- Requests from students concerned about content/stories about them in The Collegian—these requests became more frequent as we added content from 1976-1987
- The University Libraries allowed Google to crawl its database (2007/08)
- Related projects to digitize other Penn State University campus newspapers
  - Hazleton Highacres, 1937-1994
  - Behrend Beacon 1948-1989
  - Penn State Harrisburg Online Archive of Student Newspapers, 1969-2010

4 The Collegian Online website continued to evolve
An evaluation of licensed and copyrighted content—concern about the amount of copyrighted, licensed, or syndicated content published in the Post-1923 print edition resulted in the converted the post-1923 issues from article level to page-level views (See Table 2).

- Result: converted the Colle­gian Historical Digital database from article level views to page level images for post 1923 content due to concerns about copyright, licensed, and syndicated content available in the print edition (2009).
- Began an evaluation of the quality of OCR and segmentation (2004/05).

Memorandum of Understanding

In 2006, the Collegian, Inc. General Manager had granted permission to digitize issues through 1987 with the understanding that Collegian, Inc. would come to a decision about granting permission for additional future years. It seemed, based on informal conversations, that Collegian, Inc. would be little-likely to grant permission for digitizing future issues. It seemed there were two primary reasons for this: first, because they had converted post 1988 student journalism content with their own methods, they felt they could continue to do so. Second, although the Libraries had demonstrated its ability to quickly and effectively convert the print format to a searchable digital archive, Collegian, Inc. like many newspapers, wanted to serve as the sole source for all readers and researchers seeing content for those years, which were likely to be most heavily read and used by alumni and student journalists. As the strategy and transition to the Daily Collegian Online website continued to develop, it was felt the post-1988 articles in the “Collegian Archives” (on their website) would bring traffic to the website over time. However, two events opened the door for further discussion and negotiation. Collegian, Inc.:

- Upgraded its website fall 2009—this transition resulted in the loss of content from “Collegian Archives”. This content has never been recovered and the “Collegian Archives” do not provide a complete record of student journalism from 1988 to date.
- Changed General Manager and News Advisor in 2009 and 2010, respectively. (McNally, 2010)

An initial informal meeting with the new General Manager in fall 2010 explored whether she would be willing to consider future digitization projects. So again, in January 2011, a group of cautious but enthusiastic librarians met with staff from Collegian, Inc. to discuss how such a partnership would work. The group agreed: the role of the Historical Digital Collegian database would be to continue to capture the print newspaper in its entirety. There would be a delay in newspaper content appearing in the Historical Digital Collegian to ensure that the Collegian Online website served as the primary archive for current news. This approach, suggested by the new Collegian General Manager allowed the Libraries and Collegian, Inc. resulted in the following agreement:

- The next (and first conversion and load) will be years 1988-2002/03 (academic year calendar).
- Following this initial conversion The Libraries will "ingest" The Daily Collegian issues into its Digital Collegian database every three years, beginning with 2003/04.
- Future conversions will take place every three years, allowing The Daily Collegian sole access to the most recent three year’s content.

In return for allowing us to digitize issues 1988 and forward, the University Libraries agreed to:

- provide a bound copy of The Daily Collegian to The Daily Collegian, Inc. and to Special Collections Library;
- microfilm issues of The Daily Collegian for preservation and to be used for conversion to digital;
- create a searchable (by keyword) and browse-able (page by page) digital archive of the complete issues of The Daily Collegian;
- digitize, in addition to all daily issues, supplements such as Venues and other special issues (i.e., the orientation issue at the beginning of semesters; Arts Festival issues; Blue and White game issues, and Magazines for each Penn State Football games;
- provide access to the database and all content freely via its own Libraries website;
• promote and link to the Historical Digital Collegian Database and Collegian Online website.

We also were able to secure their willingness to transfer their photo archive to the University Archives and to allow us to develop a project to digitize that collection.

**Editing, Correcting, and Blocking Content: Final Negotiations**

Today’s college students live their lives in public social media spaces yet they can be surprisingly naïve and unprepared when information about them appears in a newspaper or news source. They (or their parents) can frequently be taken aback when a news story or notice catches their actions, which may not always be flattering or advantageous from the viewpoint of future employment opportunities. (Albanese, 2008) Because the Collegian Online website has been accessible since 1996, students are more likely aware that they have much less privacy should their work or lives appear on the Collegian Online website. However, because this next phase of digital conversion of the print newspaper would include 1989-1995 content, many students of that period may not fully understand that their lives were now going to be more visible and would no longer require a trip to the Library to view stories published during those years. In addition, because the libraries was digitizing the print edition three years after some of its content appeared on the Collegian Online website, there was some possibility that stories could re-appear three years later when the print edition was digitized.

Newspapers deal with, and The Daily Collegian is no exception, requests (and threats) for stories to be corrected, edited or removed on a daily basis. The Daily Collegian, like all newspapers, has always had a number of approaches to deal with these situations as they happen. (Schwartz, 1975) Some of these situations include threats of legal action which may not be actualized; in other cases, the situation a mistake has been clearly made by a student journalists and the content may need to be removed or identified for future readers. More recently, concerns of plagiarism and fabrication also haunt student run newspapers. Only rarely would content be completely removed from the website editions—even though the print edition may retain this content.

In addition, as digital archives of newspapers are created and discovered via web search engines, stories, once essentially hidden or lost in microfilm archives, may become visible once again, causing students and alumni to request stories in digital editions of newspapers be removed, edited, or updated. In addition, because of the changing nature of news delivery—The Daily Collegian is essentially running two separate news outlets—created with the same staff, and generally (at this time) including the same student-created content in both the web site and the print editions. There is, however, a great deal of sensitivity and awareness that this is a student-run newspaper reporting on university students, on the part of the staff who oversee the daily operations of the newspaper. A series of discussions related to this issue revealed two threads of concerns and values (Table 3).

Table 3: Values of a student run newspaper and library values
The Collegian valued

- Flexibility and willingness from the Libraries in considering their requests to remove content when they judged it to be appropriate to do;
- Understood the need to retain an intact archival record of The Collegian, but also were more likely to err on the side of preserving a student’s privacy or reputation;
- Had informal internal procedures for handling most of these situations, but little written documentation or public statement on their website about how such situations were appealed or addressed.

The University Libraries valued

- Wanted to retain a complete and accurate digital archival facsimile of The Daily Collegian newspaper;
- Did not want to be put in a position of being asked to remove content every time a student (or parent) threatened legal action or of having to negotiate these situations with The Collegian, Inc. staff;
- Removing content only when there was a “compelling legal reason”;
- Clear guidelines and consistent policies for how situations would be handled and encouraged The Daily Collegian to develop clear written guidelines in how these situations were handled and more transparency in providing this information to their readers.

Of course, The Collegian Online can easily remove, add, or edit content that appears on the website; changing/editing/removing content from the print edition requires a more thoughtful approach as the Collegian, Inc. maintains two editions in two different formats. With this in mind, the Libraries and Collegian, Inc. worked through several different real-life situations that had presented themselves over the past few years. These discussions resulted in developing language and ultimately working with the University’s risk management department to come an agreement on language (Table 4) which allowed the Libraries and Collegian, Inc. to retain the values most important to each.

Table 4: Types of content which may require removal, editing, or correction

<table>
<thead>
<tr>
<th>Types of corrections</th>
<th>Covered by Memorandum of Understanding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reporting that included factual errors</td>
<td>“The Libraries will consider requests to block online content from readers/researchers where libel, copyright, or other legal issues may be a consideration.”</td>
</tr>
<tr>
<td>Reporting that demonstrated a clear ethical violation</td>
<td></td>
</tr>
<tr>
<td>Reporting or legal notices which reflected poorly on the person or in which the ultimate outcome of the adjudication was not reported</td>
<td></td>
</tr>
<tr>
<td>Content that was plagiarized or fabricated</td>
<td></td>
</tr>
<tr>
<td>Content that violates copyright law or license agreements</td>
<td></td>
</tr>
</tbody>
</table>

Conclusion

The challenges of retaining a complete archival record of a student run newspaper presented the University Libraries with several challenges over a 10 year period, but ultimately the University Libraries has established a strong relationship with Collegian, Inc. and will continue to provide access to The Daily Collegian print edition through its Historical Digital Collegian database. Our next challenge will be providing access to the collection of photographs from the Collegian photographers. In addition, student newspapers (Doctor, 2012) are facing the same challenges
community and national newspapers face. (Read, 2003) The University Libraries hopes to work closely with *The Daily Collegian* as it becomes an increasingly online only newspaper and developing strategies and approaches to provide an archival record of *The Collegian Online* website.
Table 2: Evaluation of licensed, copyrighted, and syndicated content

<table>
<thead>
<tr>
<th>Date</th>
<th>Articles</th>
<th>Associated Press</th>
<th>News Briefs</th>
<th>Photographs/Graphic Illustration</th>
<th>Editorial Cartoon</th>
<th>Syndicated Cartoon Strips</th>
<th>Remaining Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>August 16, 1972</td>
<td>22</td>
<td>10</td>
<td>13</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>Advertising</td>
</tr>
<tr>
<td>August 26, 1981</td>
<td>23</td>
<td>19</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>Editorial</td>
</tr>
<tr>
<td>December 3, 1987</td>
<td>33</td>
<td>20</td>
<td>2</td>
<td>12</td>
<td>2</td>
<td>0</td>
<td>Advertising</td>
</tr>
</tbody>
</table>
**Works Cited**


The Pennsylvania State University. (2013). *Our History*. Retrieved June 10, 2013, from This is Penn State: [http://www.psu.edu/this-is-penn-state/our-history](http://www.psu.edu/this-is-penn-state/our-history)

The Pennsylvania State University. (2013a). *Rankings*. Retrieved June 10, 2013, from This is Penn State: [http://www.psu.edu/this-is-penn-state/rankings](http://www.psu.edu/this-is-penn-state/rankings)