"Reports of My Death Are Greatly Exaggerated": Findings from the TEI in Libraries Survey

Michelle Dalmau and Kevin Hawkins

"Reports of My Death Are Greatly Exaggerated": Findings from the TEI in Libraries Survey

Michelle Dalmau and Kevin Hawkins

1. Introduction

1   In the early days of the TEI Guidelines, academic libraries extended their access and preservation mandates to include electronic text (Engle 1998; Friedland 1997; Giesecke, McNeil, and Minks 2000; Nellhaus 2001). At the turn of the twenty-first century, momentum for text encoding grew in libraries as a result of the maturation of pioneering digital-library programs and XML-based web-publishing tools and systems (Bradley 2004). Libraries were not only providing "access to original source material, contextualization, and commentaries, but they also provide[ed] a set of additional resources and service[s]" equally rooted in robust technical infrastructure and noble "ethical traditions" that have critically shaped humanities pedagogy and research (Besser 2004).

2   In 2002, Suzana Sukovic posited that libraries' changing roles would and could positively impact publishing and academic research by making use of standards such as the TEI Guidelines as well as traditional library expertise, namely in cataloging departments because of their specialized knowledge in authority control, subject analysis, and bibliographic description (Sukovic 2002).

Soon thereafter, in 2004, Google announced the scanning of books in major academic libraries to be included in Google Books (Google 2012), and in 2008 many of these libraries formed HathiTrust to provide access to facsimile page images created through mass digitization efforts (Wilkin 2011). For many, the momentum behind mass digitization called into question the role for libraries in text encoding that Sukovic had advocated. In 2011, with the formation of the HathiTrust Research Center and IMLS funding of TAPAS (TEI Archiving, Publishing, and Access Service, http://www.tapasproject.org/), we see that both large- and small-scale textual analysis are equally viable and worthy pursuits for digital research inquiry in which libraries are heavily vested (Jockers and Flanders 2013).

3   More recently, we are witnessing a call for greater and more formal involvement of libraries in digital humanities endeavors and partnerships (Vandegrift 2012; Muñoz 2012) in which the resurgence of TEI in libraries is becoming apparent (Green 2013; Milewicz 2012; Tomasek 2011; Dalmau and Courtney 2011). How has advocating for such wide-ranging library objectives—from digital access and preservation to digital literacy and scholarship, from supporting "non-consumptive research" ("nonexpressive use") to supporting research practices rooted in the markup itself—informed the evolution or devolution of text-encoding projects in libraries?

4   Inspired by the papers, presentations, and discussions that resulted from the theme of the 2009 Conference and Members' Meeting of the TEI Consortium, "Text Encoding in the Era of Mass Digitization," the launch of the AccessTEI program in 2010, and the release of *Best Practices for TEI in Libraries* in 2011 (Hawkins, Dalmau, and Bauman 2011), we surveyed employees of libraries between November 2012 and January 2013 to learn more about text-encoding practices and gauge current attitudes toward text encoding in libraries. We hypothesized that as library services evolve to promote varied modes of scholarly communication and accompanying services, and as digital library initiatives become more widespread and increasingly decentralized, text encoding is undertaken less often in libraries, especially at smaller institutions, and is seeing decreased support even at larger institutions. We also wanted to investigate the nature of library-led or -partnered electronic text projects, including whether there is an increase or decrease in local mass digitization or scholarly encoding initiatives.

2. Method

5   We developed a survey using SurveyMonkey with a combination of yes/no, multiple-choice, ranking and rating, and free-response questions. In an effort to collect longitudinal data that we could leverage in our own study, we referenced and modeled a subset of questions after a survey circulated in 2008[1] that informed what we now know as AccessTEI,[2] a TEI Consortium member benefit providing a volume discount for digitization and text encoding.

6   We decided to target communities of practice as opposed to individuals, intending to lower the probability of bias that might have occurred with an otherwise judgmental sample of responses. However, we encouraged responses from multiple staff members from the same institution to ensure a more holistic view of text-encoding practices across libraries. In turn, we generated institutional biases that we did not attempt to normalize since the data was collected in an anonymous fashion.

7   We formally announced the survey as part of the poster sessions for the 2012 Digital Library Federation (DLF) Forum and the 2012 Conference and Members' Meeting of the TEI Consortium, which occurred within weeks of each other.[3] Once the survey was unveiled at the DLF Forum on November 4, 2012, the survey was announced via digital library and digital humanities mailing lists (including TEI-L, DLF-ANNOUNCE, DIGLIB, and XML4LIB) and social media channels like Twitter and Facebook.

8   Depending on how they answered certain questions, respondents encountered one of four paths with 11, 17, 28, or 30 questions to complete. The only respondent requirement was that she or he worked in a library in some capacity. Not all questions were answered; we have estimated a completion rate of 60% that takes into account the various forks in the survey.

9   The survey was comprised of four major sections:

- Study Information
- Determination of Eligibility
- Background Information About the Institution (type of library, size, attitudes)
- Text Encoding Practices

    ○   Standards Used

    ○   Collaborations/Partnerships

      ◦   Types of Text Encoding Projects

10   The survey closed on January 31, 2013, with 112 valid responses that provided the foundation for our analysis. The survey questions and the data collected are available at https://github.com/mdalmau/tei_libraries.

2.1 Data Preparation

11   Mishaps occurred with the data collection using SurveyMonkey because of a combination of researcher error and glitches with the survey tool. Of the original 138 respondents, 26 answered "no" to the question "Do you work in a library?," but despite their not meeting the sole criterion for taking the survey, the system somehow allowed them to continue. These responses were disqualified. In addition, a subset of questions (for 10 respondents) were disqualified (and marked as "invalid") based on other errors uncovered in SurveyMonkey's logic for skipping questions.[4] The Indiana Statistical Consulting Center provided close consultation to determine which responses to disqualify and to normalize the data for statistical processing, which included content analysis and coding of the qualitative responses.

12   We coded responses of both the quantitative and qualitative questions. After questions were keyed (Q1, Q2, etc.) for statistical processing, values for all ranking questions, Likert scale questions (with responses ranging from "almost always" to "never"), and yes/no questions were normalized. Six qualitative questions (Q4, Q9, Q16, Q25, Q118, and Q119) were coded following a three-step process: (1) each of us coded the responses separately, (2) we combined our respective codings to generate a single scheme, and (3) together we reassigned codes based on the single scheme.

13   The spreadsheet containing the coded data, which is also available on GitHub (https://github.com/mdalmau/tei_libraries), contains multiple tabs including

- Q_KEY, containing the mapping of the prose questions to an identifier scheme for statistical processing
- Data with the 112 valid responses including normalized values
- Likert_Key, reflecting the normalized values assigned to all Likert scale questions
- Content analysis of the 6 qualitative questions including original and coded responses:

◦   Q4, Q9, Q16, Q25, Q118, Q119 (each in separate tabs).

3. Results

14   The following summary and discussions of the results are presented as a "snapshot" in time based on analysis of data collected in the survey. The lack of pre-existing data measuring text-encoding activities in libraries made it difficult to make assertions about changes over time. Still, the results provide valuable information about text-encoding activities and attitudes in libraries that can be used in future studies.

3.1 Profile of Survey Respondents

15   Of the 112 respondents, we determined from IP addresses that

- 55 are clearly affiliated with an institution; 41 of these are unique institutions
- 57 are unidentifiable because they used off-site internet connections (via ISPs).

16   As table 1 indicates, most respondents are affiliated with North American academic libraries. This finding is not surprising given that publicity was mostly on North American and UK lists and given the relatively long history of American academic library support and adoption of the TEI Guidelines, starting with the TEI and XML in Digital Libraries Workshop sponsored by the Digital Library Federation (DLF) in 1998 (Hawkins, Dalmau, and Bauman 2011). In 1999, the DLF published the first version of what was known as the "TEI in Libraries Guidelines" (Digital Library Federation 1999), and in 2011, version 3, now known as *Best Practices for TEI in Libraries* (Hawkins, Dalmau, and Bauman 2011), was released, with contributions mostly by American librarians.

Table 1. Responses to demographic questions pertaining to the respondent's institutional affiliation (n=112).

Where is your institution located?	Indicate the type of library for which you work.	What is the size of your academic institution based on student enrollment or number of patrons served?

Asia	2	Academic Library	92	Up to 5,000	8
Europe	11	National Library	2	5,000–10,000	16
North America	84	Public Library	5	10,000–25,000	27
Unknown	15	Research Library	10	25,000–40,000	22
No Response	0	Special Library	3	Over 40,000	14
		No Response	0	No Response	25

17   Respondents were asked to identify their departmental affiliations, and to list departments with which they partner on text-encoding projects. Responses were coded (see figure 1) according to twelve main areas of work or departments (such as cataloging or technology), but not weighted with respect to respondents providing multiple departmental affiliations (9 of 112 responses). Not surprising, departments reporting the most text-encoding work include Technology, Digital Scholarship, Cataloging, Special Collections, and Archives. Of the 58 respondents who indicated units with which they partner, most had partnered with at least 3 other departments elsewhere in the library, revealing a concentration of partnerships in departments like Technology, Digitization, and Cataloging. While we cannot claim that text-encoding work has become "decentralized" in libraries based on our data alone, we certainly see a spread of text-encoding work across various library departments (see figure 1).

Figure 1. This pie chart shows respondents' reported departmental affiliations, coded according to twelve main areas of work or departments, with "General Library" for responses such as "main" and "general."
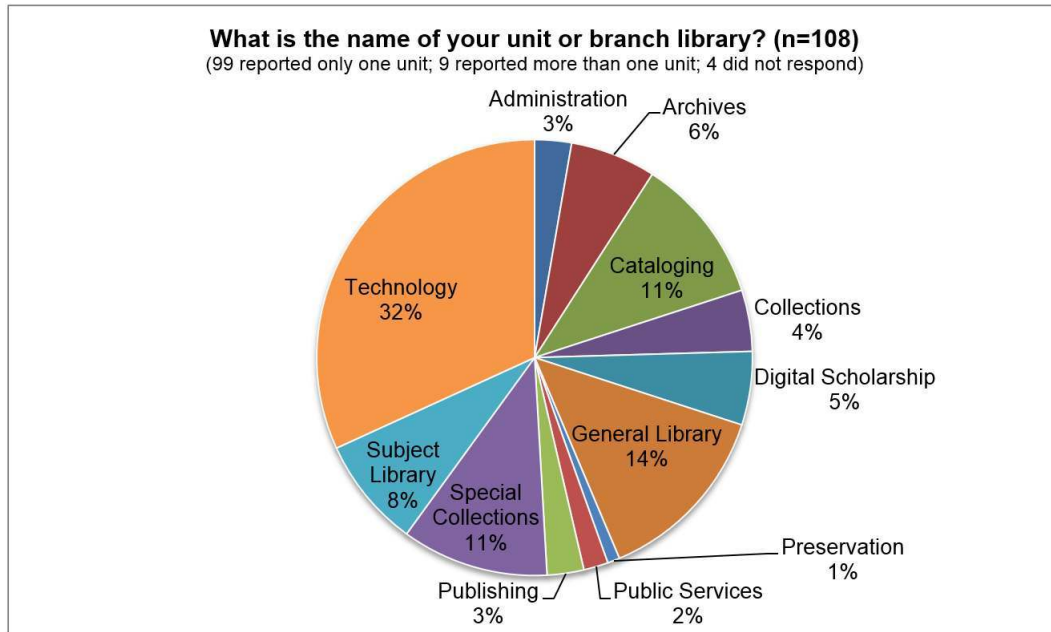
3.2 TEI Consortium Affiliations

18 As mentioned earlier, we were primarily interested in the individual's experience with text-encoding practices in libraries, but we also asked respondents to identify whether their institution was affiliated with the TEI Consortium. To present a more accurate picture, we attempted, only in this instance, to control for multiple responses per institution (figure 2). This chart gives four data points for each response:

- total responses
- total institutions
- total unique institutions
- total respondents who accessed the survey via an ISP.

19 The responses were analyzed based on whether or not the respondent said his or her institution is a member of the TEI Consortium, in addition to unsure and blank responses. For those who answered "yes" to the TEI Consortium membership question, we can see that 18 of the 39

respondents are affiliated with an identifiable institution and 9 of those (half), after de-duplication, are unique institutions. For those who answered "no" to the question, we can see that 23 of the 43 respondents are affiliated with an identifiable institution and 17 of those are unique institutions. In sum, 50% of respondents that claimed their institutions are members of the TEI Consortium were identified as being from unique institutions, and 73% that claimed their institutions were not affiliated were identified as being from unique institutions. In keeping with Lynne Siemens' report, "Understanding the TEI-C Community: A Study in Breadth and Depth, Toward Membership and Recruitment," presented at the TEI Consortium's Members Business Meeting in 2008, [5] it is not surprising that most respondents are not from institutions that are members of the TEI Consortium. The report, which led to the publication of "The Apex of Hipster XML GeekDOM" (Siemens et al. 2011) covered the findings of a study led by Lynne Siemens and her team to survey the at-large TEI community's response to a viral video featuring the TEI encoding of Bob Dylan's "Subterranean Homesick Blues" in order to understand current TEI membership and potential members.

Figure 2. This graph shows TEI Consortium membership status as reported by respondents, with an attempt to de-duplicate institutional affiliations as indicated by the "Total Unique Institutions" data point.
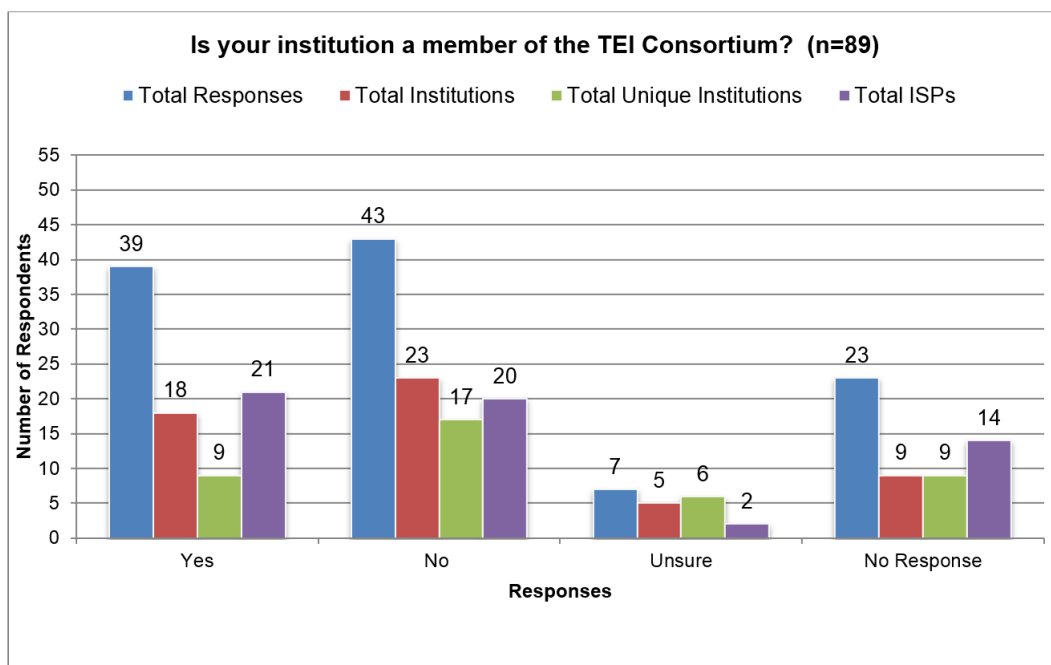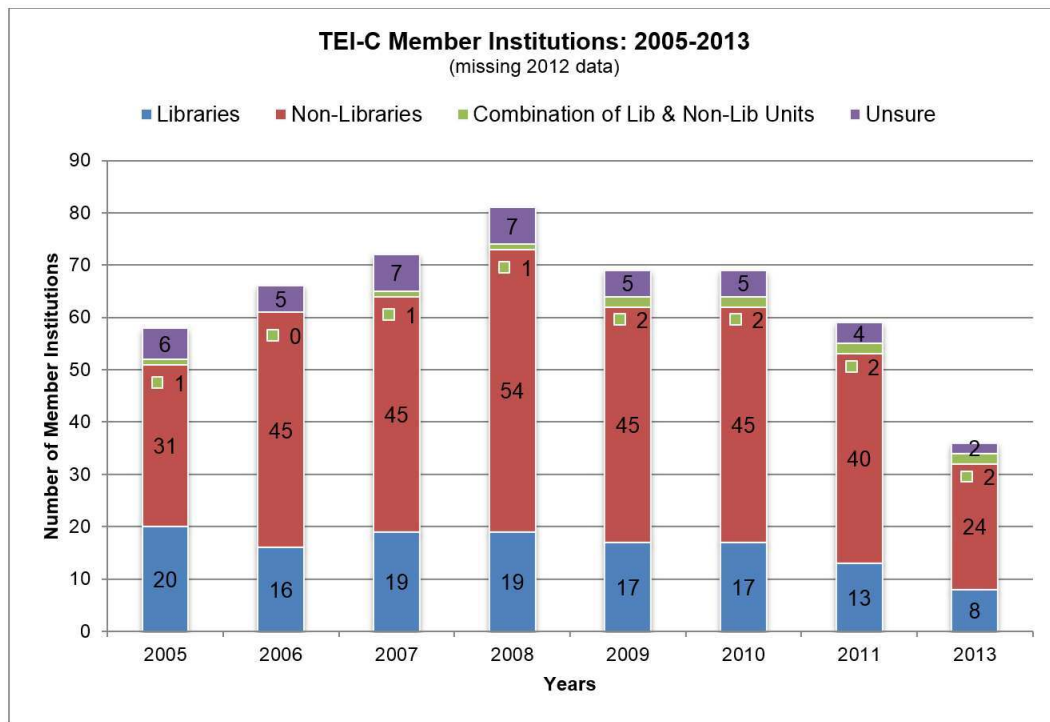
Figure 3. This graph shows the number of TEI-C member institutions from 2005 to 2013 (with the exception of 2012) coded by type: libraries, non-libraries, combination of library and non-library units, and unsure. The 2012 data is incomplete due to a change in the TEI-C membership tracking system.

20   We attempted to compare the TEI Consortium membership data we collected with the TEI's historical membership records (for 2005 to 2011 and 2013). We coded the institutions as one of the following: libraries, non-libraries, a combination (as represented by partnering units like an academic or technology department and a library), or unsure (see figure 3).

21   We see a fairly consistent membership base consisting of an average of 18 member libraries between 2005 and 2010. If we correlate membership data with the start and rise of mass digitization like Google Books (2004) and HathiTrust (2008), we do not see an obvious impact of these initiatives on library membership. The decline in membership we do see in 2011 and 2013 is not unique to library members and could very well be an effect of the 2008/2009 global recession, which led to significant reductions in library budgets (Valade-DeMelo 2009; Bailey 2009; Nicholas et al. 2010).

22   Still, simply counting member institutions does not reflect the varying level of financial support that they offer to the TEI through different classes of membership. While it is often said that libraries provide the majority of financial support for the TEI Consortium, it turns out that library members contribute an average of 45% of the TEI-C's revenue (Hawkins 2014); not quite half, but indeed a significant collective contribution.

3.3 Text-Encoding Practices and Partnerships in Libraries

23   Libraries support text encoding across a wide spectrum of discrete tasks and work practices associated with starting and completing a text-encoding project, from consulting and training to actual markup and web publishing (see figure 4). Such activities are carried out in partnership with various other constituencies inside and outside the library (see figure 5). As we have seen thus far, it is not surprising that the greater number of partnerships is across library staff and departments, but we see an equally high number of partnerships with faculty and information technology (IT) staff. Such library-faculty partnerships could indicate a trend toward more advanced or scholarly text-encoding support. How tasks align with partnerships is not surprising: for example, we see IT staff featuring prominently in web-publishing tasks, and librarians featuring prominently in establishing text-encoding workflows and engaging in markup directly. Despite the nature of the relationship, what is of particular interest is the great number of faculty partnering with libraries on text-encoding projects.

Figure 4. This graph shows ways in which respondents reported that they support text-encoding activities in their respective units.
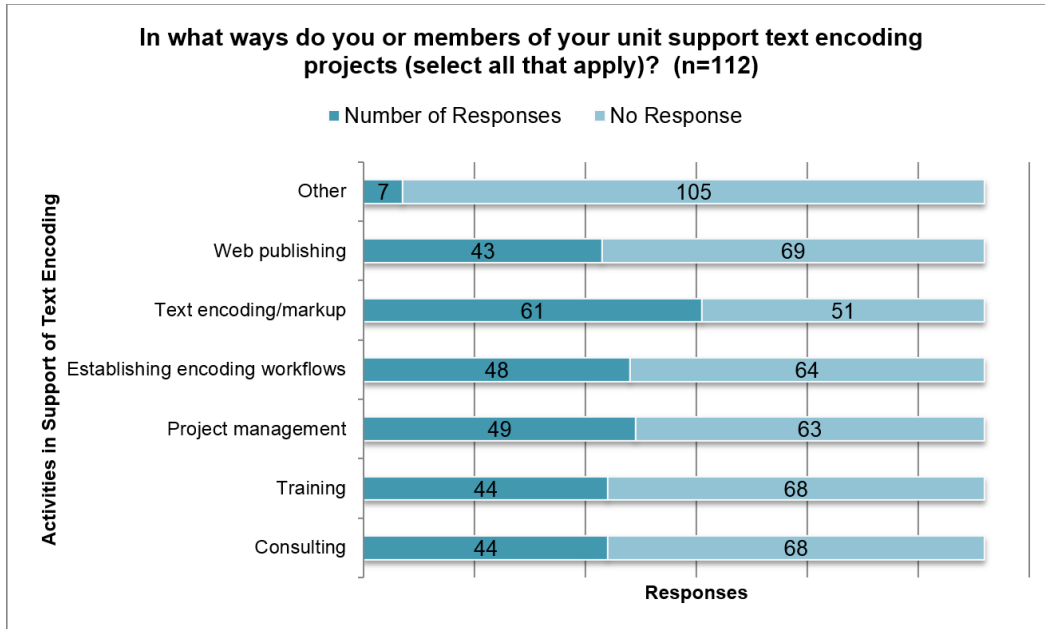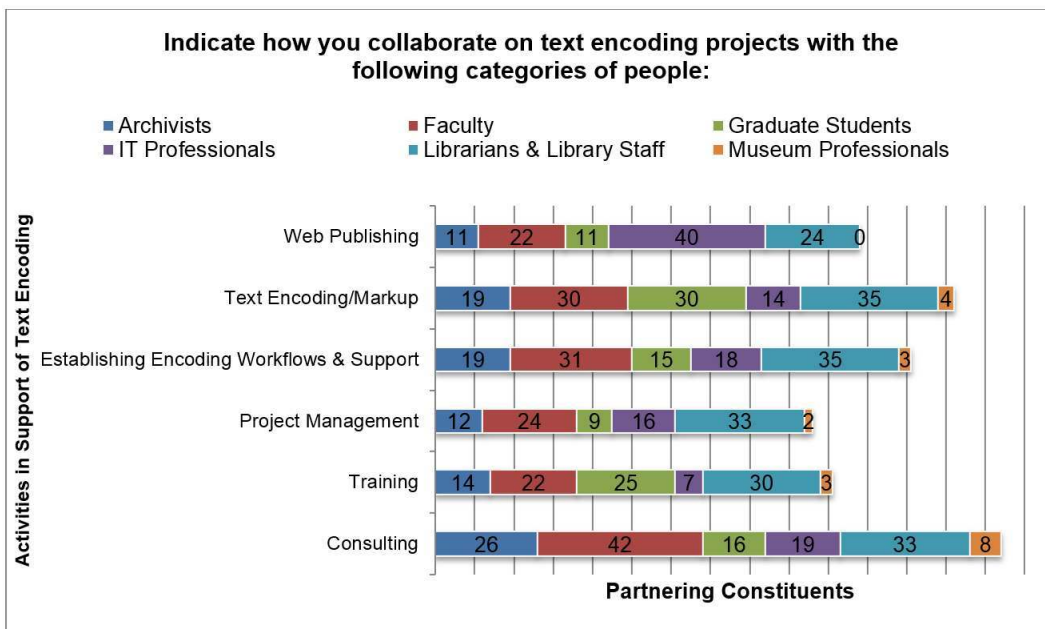

Figure 5. This graph shows ways in which respondents reported partnering with other constituencies on text encoding.

24 Respondents were asked to rank eight types of projects or kinds of collections commonly encountered in libraries in terms of how often they work with such collections, from most common to least common. As is evident in figure 6, based on the data reflected in table 2, the top three most common types of projects or collections for which text encoding features prominently are rare books and manuscripts, archival materials, and faculty or librarian digital research projects. It appears that text encoding is reserved for special collections and unique content, not the most commonly used materials.

Figure 6. This graph shows the frequency of the three most common responses to the question "Rank the nature of your text encoding projects (1 is most common, 8 is least common)": Rare Books & Manuscripts, Archival Materials, and Faculty or Librarian Digital Research Projects.
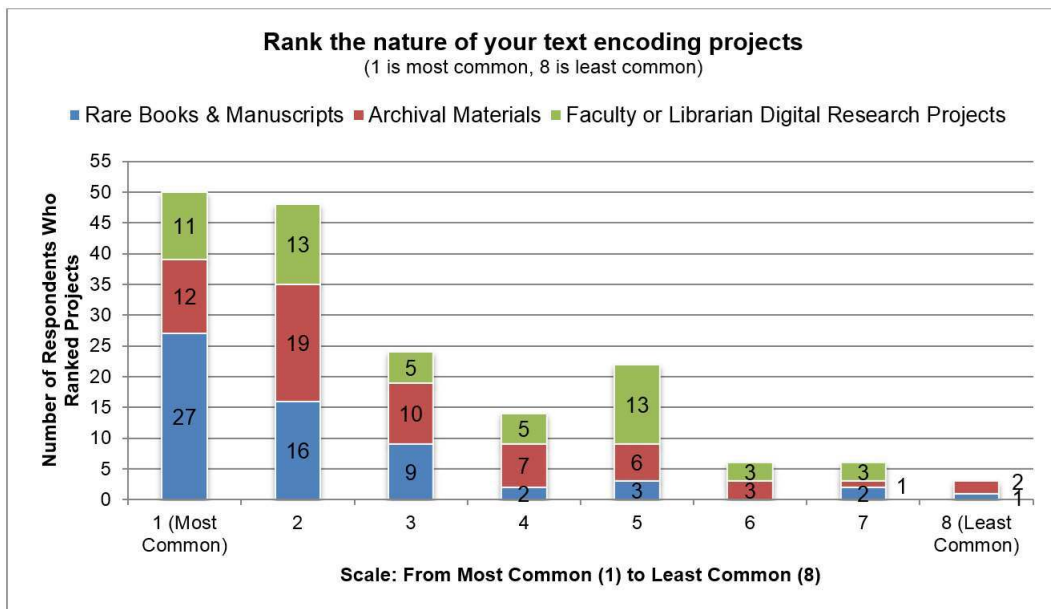

Table 2. This table shows all responses to the question "Rank the nature of your text encoding projects (1 is most common, 8 is least common)."

	1 (Most Common)	2	3	4	5	6	7	8 (Least Common)	N/A or Blanks
Archival Materials	12	19	10	7	6	3	1	2	60

Faculty or Graduate Student Digital Teaching Projects	2		5	5	5	6	15	6	7		61
Faculty or Librarian Digital Research Projects	11		13	5	5	13	3	3	0		59
Library General Collections	4		3	4	11	14	5	4	2		65
Other Library Special Collections	3		6	17	21	2	3	3	1		56
Rare Books & Manuscripts	27		16	9	2	3	0	2	1		52
University Collections	1		0	5	1	7	9	16	1		72
University Press	7		1	0	0	0	6	6	1		70

25   We asked respondents to describe the level of text encoding with which they most often engage, describing these levels abstractly rather than as numbers as in the *Best Practices for TEI in Libraries:*

- Basic reformatting of text for bibliographic and keyword search (Level 1)
- Mid-level structural encoding for full-text display and basic functionality like linking table of contents, notes, etc. (Levels 2 and 3)
- Richer encoding for content analyses like name tagging, rhyme schemes, etc. (Level 4)
- Scholarly encoding projects (Level 5)

Figure 7. This graph shows the frequency that respondents reported conducting different types of encoding.
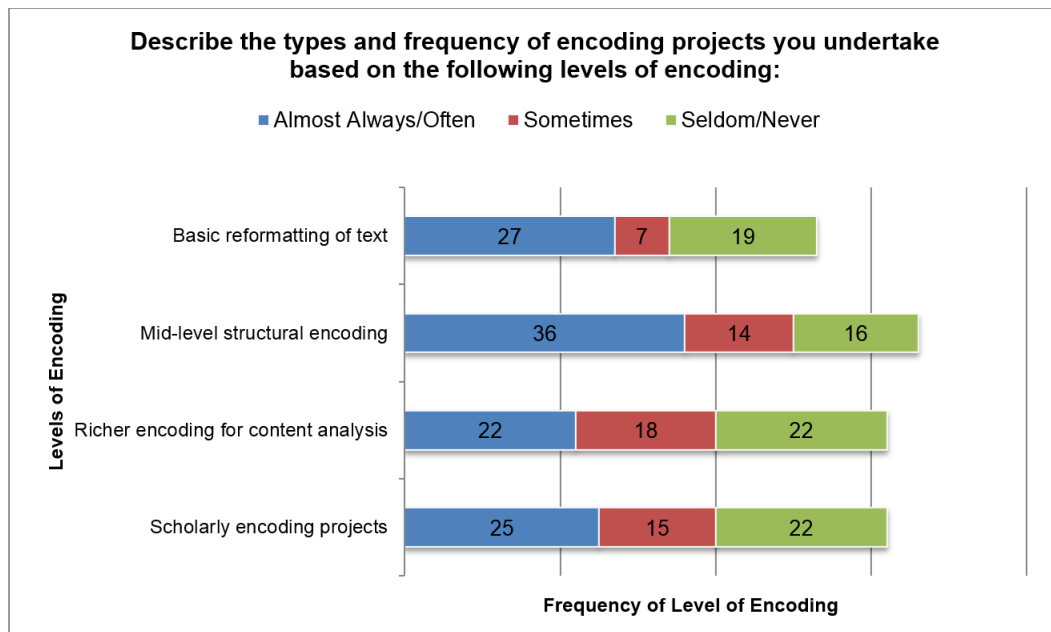
26   According to figure 7, we can see activity across all levels of text encoding with an emphasis on mid-level structural encoding. We also asked respondents to indicate the number of text-encoding projects with which they are involved, from none to more than 30, with most people working on 1–5, 6–10, or more than 30 projects. We correlated the number of projects with encoding levels, and assumed that those involved with fewer projects are encoding at higher levels and vice versa. Instead, we noticed a wide range of activity across all levels of encoding regardless of the number of text-encoding projects. However, as we look more closely at the correlation between levels of encoding and types of materials most commonly encoded in libraries (figure 8), we see peaks in mid-level structural encoding (level 3), richer encoding for content analysis (level 4), and scholarly encoding (level 5).

Figure 8. This graph shows the frequency of different types of encoding for two types of material reported as the most commonly encoded.
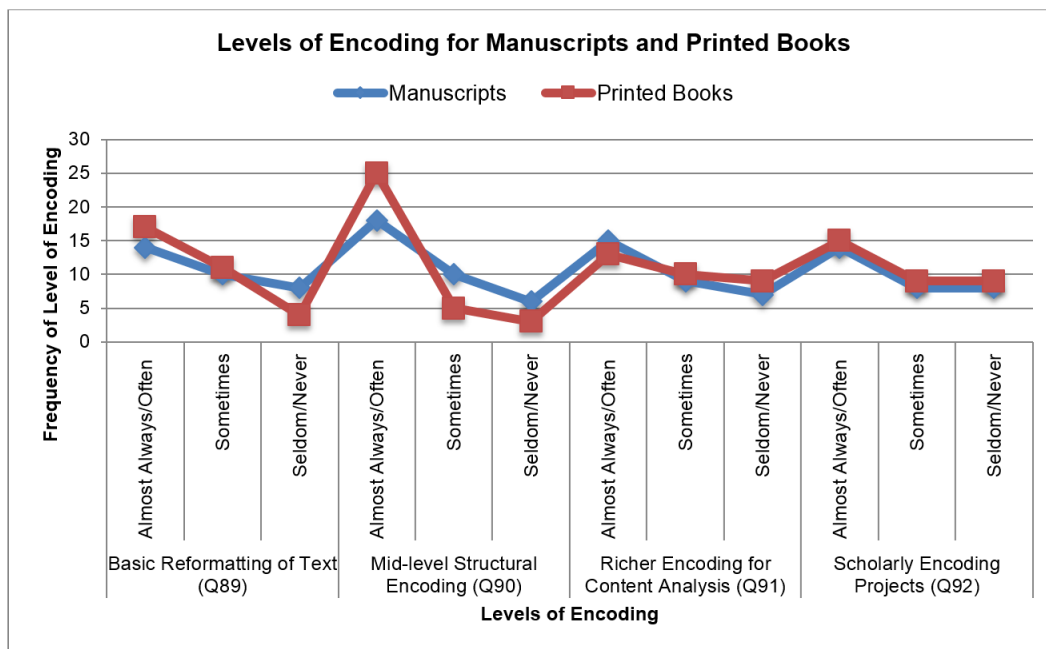

3.4 Text-Encoding Interests and Attitudes

27　We presented respondents with a mixture of quantitative and qualitative questions with respect to text-encoding interests and attitudes across their library. We correlated responses to both sets of questions to ensure reliability of the responses.

Table 3. This chart shows responses (n=112) to survey questions in which respondents rated their library's administrative support for text-encoding projects and general level of interest in text-encoding projects across the library as a whole.

How would you rate your library's administrative support for text encoding projects?		How would you rate the level of interest in text encoding by members of your library as whole?	
Extremely Supportive	5	Extremely Interested	1
Very Supportive	21	Very Interested	10
Moderately Supportive	37	Moderately Interested	36

Slightly Supportive	23	Slightly Interested	39
Not at all Supportive	13	Not at all Interested	16
Not Applicable	6	Not Applicable	4
No Response	7	No Response	6

Figure 9. This graph shows a cross-tabulation of reported administrative support for text encoding and reported general interest across the respondent's library in text encoding.
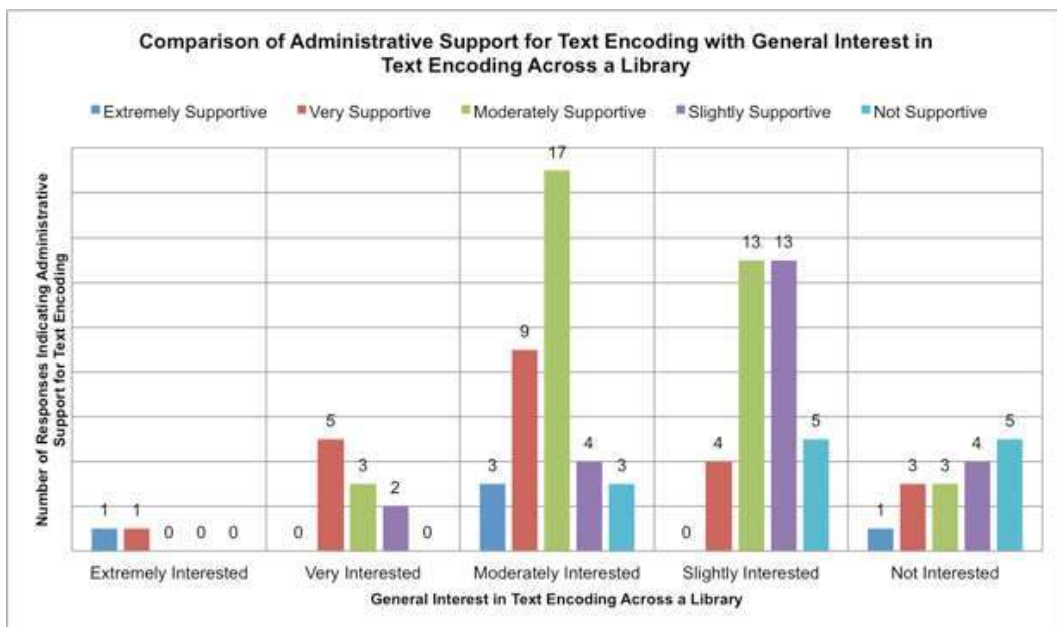
28 As seen in table 3 and figure 9, administrative support and general interest in text encoding across the libraries are closely related, as they are respectively situated in the moderately-to-slightly-interested and moderately-and-slightly-supportive responses of the Likert scale.

29 At face value, this occupation of the middle ground might seem like a safe, even rational, place for an institution given this time of transition for academic libraries as they begin to define themselves more clearly in the age of digital scholarship. However, the sentiments on the fringes of the Likert

scale are problematic. We see little to no correlation between an extremely and very supportive library administration and an extremely and very interested library staff. And the "not interested" camp is threatening to tip the moderate scale.
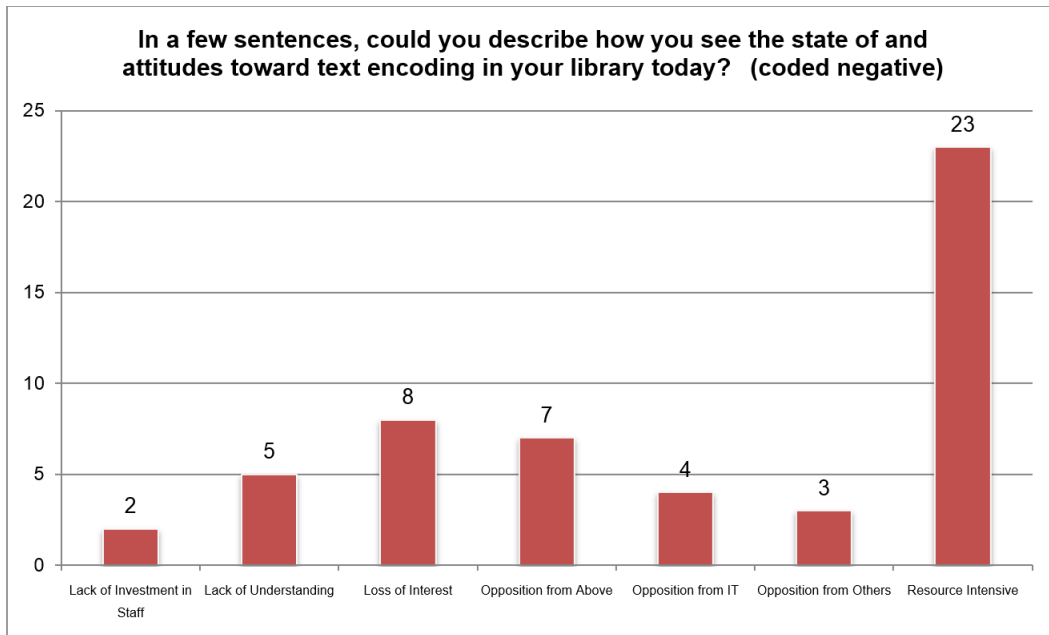
30   We then compared the quantitative responses to qualitative responses we collected. A little over half of the respondents (approximately 63 of 112) answered the question: "In a few sentences, could you describe how you see the state of and attitudes toward text encoding in your library today?" We completed two levels of coding for the qualitative responses to this question: we assigned thematic categories to the responses (following the three-step process identified above in the "Data Preparation" section), and then we tagged the categories as either positive, negative, or neutral. For this analysis, we did not disqualify those who only provided responses to the quantitative questions, though we disclose the number of "no responses" (). Because of this discrepancy, the quantitative responses are marginally inflated, but they do not seem to detract or bias the qualitative responses in any way, as is made clear by their strong correlation.

31   Those in the neutral camp (35%) align well enough with the slightly-to-moderately-interested/supportive camp as seen in figure 9. The negative responses dominate at 44%, which illustrates a perceived threat to text encoding in libraries (see figure 11), leaving 21% positive responses.

32   Figure 10 reveals the categories coded as positive and their distribution among respondents. The granular coding makes it impossible to generalize these sentiments more broadly, but the number of people who reported "expected uptake" and "general interest" in text-encoding projects is heartening. Responses indicating that the survival of text encoding in their library is a result of individual initiative are more concerning since this implies an overall lack of institutional support. Though the numbers are not as high, interest among catalogers and the training opportunities around text encoding correlate with trends we are seeing in figures 1 and 4.

Figure 10. Of responses (n=63) to the question "In a few sentences, could you describe how you see the state of and attitudes toward text encoding in your library today?" this graph shows responses with portions coded as positive (n=25) after two levels of coding: (1) themes were identified and then (2) themes were tagged as positive, negative, or neutral.
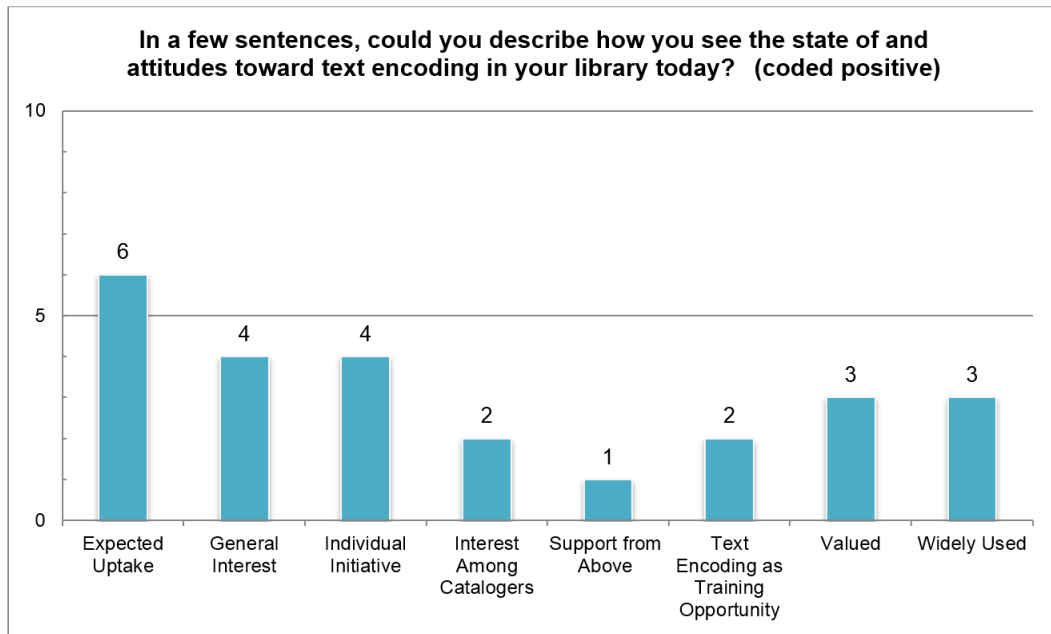
33   The findings for the categories coded as negative are not especially surprising (figure 11). Libraries have been struggling with the resource intensity of text encoding, from doing markup to publishing the encoded texts online, for years. The various types of opposition to text encoding reported require further exploration. While we have not correlated the responses coded as various types of opposition with responses indicating that text encoding is resource-intensive, we suspect a tight relationship.

Figure 11. Of responses (n=63) to the question "In a few sentences, could you describe how you see the state of and attitudes toward text encoding in your library today?" this graph shows responses with portions coded as negative (n=52) after two levels of coding: (1) themes were identified and then (2) themes were tagged as positive, negative, or neutral.

In a few sentences, could you describe how you see the state of and attitudes toward text encoding in your library today?   (coded negative)

Theme	Count
Lack of Investment in Staff	2
Lack of Understanding	5
Loss of Interest	8
Opposition from Above	7
Opposition from IT	4
Opposition from Others	3
Resource Intensive	23

34   The neutral camp yielded a medley of codings (figure 12), including expressions of apathy, mixed feelings about whether text encoding is a viable endeavor for libraries, and uncertainty about the benefits of encoding. A few codings, however, were used for ambiguous responses that could easily manifest as positive or negative depending on the argument made. However, two themes emerged which can be seen as somewhat at odds: libraries engage selectively in text encoding, yet they prioritize basic access to text collections. Specifically, while we know that text encoding is more often used for special collections and for scholarly projects than for general collections, libraries are under pressure to provide online access to their general collections as well, forgoing encoding for simple digitization using facsimile page images and keyword searching.

Figure 12. Of responses (n=63) to the question "In a few sentences, could you describe how you see the state of and attitudes toward text encoding in your library today?" this graph shows responses with portions coded as neutral (n=42) after two levels of coding: (1) themes were identified and then (2) themes were tagged as positive, negative or neutral.

In a few sentences, could you describe how you see the state of and attitudes toward text encoding in your library today?   (coded neutral)
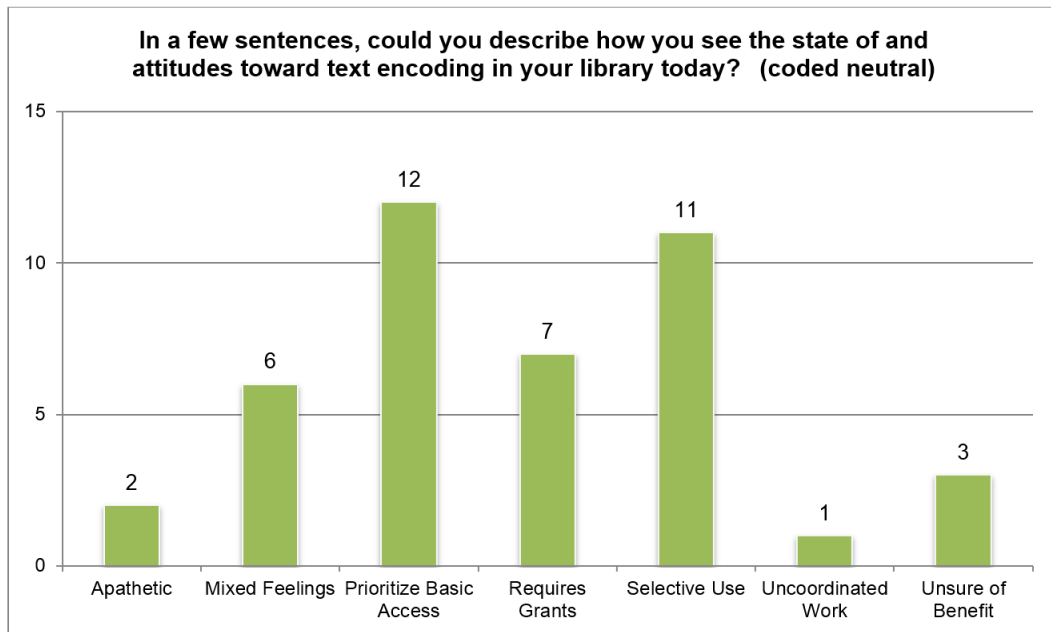
4. Discussion

35 We have uncovered several areas that deserve additional investigation and consideration. As we move forward, the "TEI and libraries" community would benefit from

- gaining a more global perspective and understanding of text encoding in libraries

- proposing TEI Consortium member benefits for libraries of all sizes with a special emphasis on cohesive, centralized, and certified training opportunities offered by the Consortium

- verifying what appears to be a concerted effort by libraries to use text encoding for special collections, and determining to what extent that correlates with the peaks we observed in structural encoding (level 3), richer encoding for content analysis (level 4), and scholarly encoding (level 5). In understanding the nature of these collections and scenarios in which text encoding is deemed important for discovery of these collections, we would be better positioned to provide fine-tuned, relevant training, guidelines, and overall support for libraries.

- exploring ways in which text encoding is resource intensive, focusing both on easing the publishing process for libraries and on libraries facilitating ways in which scholars can self-publish. These options might include: better promotion of *Best Practices for TEI in Libraries*,which now contains schemas for encoding at levels 1 through 4; understanding how libraries can benefit from and contribute to the TAPAS project; and supporting TEI Simple (formerly "TEI-Nudge," see Mueller 2013). These three initiatives imply a strong role libraries can take, with the TEI Consortium's help, in fostering TEI-aware publishing systems.

36 The limitations of the survey and the lack of longitudinal data temper any conclusions that could be drawn from the survey results. Conveyed herein is at most a snapshot of TEI in libraries today, but a snapshot with great promise. This study dovetails with more recent research conducted by Harriett Green (2012, 2013) that aimed to identify concrete ways in which libraries can foster and support text encoding for library and scholarly research projects. Though we have yet to consult these and other related data sources systematically, we have released our own data set for others to use.

37 In retrospect, we consider this survey to be a preliminary data-gathering instrument. The findings as summarized above debunk our wholesale hypothesis that text-encoding practices have significantly declined in libraries. However, the data we have gathered alone are not robust enough to make more specific claims about the state of text encoding in libraries. We are more acutely aware of this precarious "middle zone" of neither giving up on nor fully embracing text encoding that libraries are occupying and will focus our investigations on uncovering and understanding the nuances of being in the middle as a way to further refine this study.

BIBLIOGRAPHY

Bailey, Charles W., Jr. 2009. "Seven ARL Libraries Face Major Planned or Potential Budget Cuts." *DigitalKoans* (blog), April 28. http://digital-scholarship.org/digitalkoans/2009/04/28/seven-arl-libraries-face-major-planned-or-potential-budget-cuts/.

Besser, Howard. 2004. "The Past, Present, and Future of Digital Libraries." In *A Companion to Digital Humanities*, edited by Susan Schreibman, Ray Siemens, and John Unsworth. Oxford: Blackwell. http://www.digitalhumanities.org/companion/.

Bradley, John. 2004. "Text Tools." In *A Companion to Digital Humanities*, edited by Susan Schreibman, Ray Siemens, and John Unsworth. Oxford: Blackwell. http://www.digitalhumanities.org/companion/.

Dalmau, Michelle, and Angela Courtney. 2011. "The Victorian Women Writers Project Resurrected: A Case Study in Sustainability." Paper presented at Digital Humanities 2011: Big Tent Humanities, Palo Alto, California, June 19–22.

Digital Library Federation. 1999. *TEI Text Encoding in Libraries: Guidelines for Best Encoding Practices*, version 1.0. http://old.diglib.org/standards/tei-old.htm.

Engle, Michael. 1998. "The Social Position of Electronic Text Centers." *Library Hi Tech* 16(3/4): 15–20, 42. http://dx.doi.org/10.1108/07378839810304522.

Friedland, LeeEllen. 1997. "Do Digital Libraries Need the TEI? A View from the Trenches." Paper presented at TEI10: The Text Encoding Initiative Tenth Anniversary User Conference, Providence, Rhode Island, November 14–16. http://www.stg.brown.edu/conferences/tei10/tei10.papers/friedland.html.

Giesecke, Joan, Beth McNeil, and Gina L. B. Minks. 2000. "Electronic Text Centers: Creating Research Collections on a Limited Budget, The Nebraska Experience." *Journal of Library Administration* 31(2): 77–92. http://digitalcommons.unl.edu/libraryscience/63/.

Google. 2012. "Google Books History." Last modified December 21. http://www.google.com/googlebooks/about/history.html.

Green, Harriett. 2012. "Library Support for the TEI: Tutorials, Teaching, and Tools." Paper presented at TEI and the C(r|l)o(w|u)d: 2012 Annual Conference and Members' Meeting of the TEI Consortium, College Station, Texas, November 8–10.

———. 2013. "TEI and Libraries: New Avenues for Digital Literacy?" *dh+lib: Where Digital Humanities and Librarianship Meet*, January 22. http://acrl.ala.org/dh/2013/01/22/tei-and-libraries-new-avenues-for-digital-literacy/.

Hawkins, Kevin. 2014. "On the Relative Weight of Libraries among Members of the TEI Consortium." *Ultra Slavonic* (blog), January 12. http://www.ultraslavonic.info/blog/?p=46.

Hawkins, Kevin, Michelle Dalmau, and Syd Bauman. 2011. *Best Practices for TEI in Libraries*, version 3.0. http://purl.oclc.org/NET/teiinlibraries.

Jockers, Matthew L., and Julia Flanders. 2013. "A Matter of Scale." Keynote presentation at Boston-Area Days of DH 2013, Boston, MA, March 18. http://digitalcommons.unl.edu/englishfacpubs/106/.

Milewicz, Liz. 2012. "Why TEI? Text > Data Thursday." *Duke University Libraries News, Events, and Exhibits* (blog), September 26. http://blogs.library.duke.edu/blog/2012/09/26/why-tei-text-data-thursday/.

Mueller, Martin. 2013. "TEI-Nudge or Libraries and the TEI." *Blog of the Center for Scholarly Communication & Digital Curation*, October 1. http://sites.northwestern.edu/cscdc/?p=872.

Muñoz, Trevor. 2012. "Digital Humanities in the Library Isn't a Service." *Notebook* (blog), August 19. http://trevormunoz.com/notebook/2012/08/19/doing-dh-in-the-library.html.

Nellhaus, Tobin. 2001. "XML, TEI, and Digital Libraries in the Humanities." *portal: Libraries and the Academy* 1(3): 257–77. http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v001/1.3nellhaus.html. doi:10.1353/pla.2001.0047.

Nicholas, David, Ian Rowlands, Michael Jubb, and Hamid R. Jamali. 2010. "The Impact of the Economic Downturn on Libraries: With Special Reference to University Libraries." *The Journal of Academic Librarianship* 36(5): 376–82. doi:10.1016/j.acalib.2010.06.001.

Siemens, Lynne et al. 2011. "The Apex of Hipster XML GeekDOM." *Journal of the Text Encoding Initiative* (1). http://jtei.revues.org/210. doi: 10.4000/jtei.210.

Sukovic, Suzana. 2002. "Beyond the Scriptorium: The Role of the Library in Text Encoding." *D-Lib Magazine* 8(1). http://www.dlib.org/dlib/january02/sukovic/01sukovic.html.

Tomasek, Kathryn. 2011. "Digital Humanities, Libraries, and Scholarly Communication." *Doing History Digitally* (blog), November 2. http://kathryntomasek.wordpress.com/2011/11/02/digital-humanities-libraries-and-scholarly-communication/.

Valade-DeMelo, Lisa. 2009. *Impact of the Economic Recession on University Library and IT Services*. London: Joint Information Systems Committee. http://www.webarchive.org.uk/wayback/archive/20140614203644/http://www.jisc.ac.uk/publications/research/2009/libsitimpacts.aspx.

Vandegrift, Micah. 2012. "What is Digital Humanities and What's It Doing in the Library?" *In the Library with the Lead Pipe* (June 27). http://www.inthelibrarywiththeleadpipe.org/2012/dhandthelib/.

Wilkin, John. 2011. "HathiTrust's Past, Present, and Future." Remarks presented at the HathiTrust Constitutional Convention, Washington, DC, October 8. http://www.hathitrust.org/blogs/perspectives-from-hathitrust/hathitrust039s-past-present-and-future.

NOTES

1 Daniel Paul O'Donnell to TEI-L mailing list, June 19, 2008, http://listserv.brown.edu/archives/cgi-bin/wa?A2=TEI-L;MCZ7mw;200806191855060600.

2 "AccessTEI: A Digitization Benefit for Members of the Text Encoding Initiative,"last updated June 14, 2010, http://www.tei-c.org/AccessTEI/.

3 "Do You TEI? A Survey of Text Encoding Practices in Libraries" (poster presented at the 2012 Digital Library Federation Forum, Denver, CO, November 4), http://www.slideshare.net/kshawkin/20121104-14988094.

4 For more information about the data set, and access to the raw data and survey questions, visit https://github.com/mdalmau/tei_libraries.

5 "Understanding the TEI-C Community' Study," TEI Members Business Meeting, November 7, 2008: Minutes, last updated November 7, 2008, http://www.tei-c.org/Membership/Meetings/2008/mm48.xml.

ABSTRACT

In the early days of the TEI Guidelines, academic libraries extended their access and preservation mandates to include electronic text, providing expertise in authority control, subject analysis, and bibliographic description. But the advent of mass digitization efforts involving simple scanning of pages and OCR called into question such a role for libraries in text encoding. This paper presents the results of a survey targeting library employees to learn more about text-encoding practices and to gauge current attitudes toward text encoding.

INDEX

AUTHORS

MICHELLE DALMAU

Michelle Dalmau, Head of Digital Collections Services at the Indiana University Libraries, manages digital projects and services for the Libraries and affiliated cultural heritage organizations, leads e-text initiatives at the IU Libraries, and partners with faculty on scholarly text-encoding projects like Chymistry of Isaac Newton and Swinburne Project. She co-convened the TEI Libraries SIG from 2007 to 2013 and co-edited Best Practices for TEI in Libraries. She currently leads a praxis-based professional development initiative, "Research Now: Cross-Training for Digital Scholarship," for the IU Libraries' Scholars' Commons. Her undergraduate background is in English and Art History; she holds MLS and MIS degrees from Indiana University.

KEVIN HAWKINS

Kevin S. Hawkins is director of library publishing for the University of North Texas Libraries. Previously he was director of publishing operations for Michigan Publishing, the hub of scholarly publishing at the University of Michigan Library which includes the University of Michigan Press and other brands and services. Kevin has also worked as visiting metadata manager for the Digital Humanities Observatory, a project of the Royal Irish Academy. His involvement with the TEI includes co-editing the 2011 revision to *Best Practices for TEI in Libraries* and serving for two terms on the Technical Council.