Dawn Schmitz, Ph.D., MLIS
Archivist
University of California, Irvine
Research focus: Media History and Cultural Studies

SUMMARY

1. Introduction

This report analyzes three online resources for discovering and using mass-digitized book content: Google Book Search (GBS), Microsoft Live Search Books (MLSB), and the Internet Archive's (IA), text search. The purpose of the investigation reported on here was to test the utility of these resources for scholarly purposes, and in particular for carrying out research in media and cultural studies. Starting with a brief bibliography of key texts in three disciplinary areas and using the three online resources for further discovery of relevant texts, I analyzed 88 different volumes, approximately 40 percent from GBS, 40 percent from MLSB, and 20 percent from IA. I viewed up to 100 pages of text for each volume (depending upon how many were made available) to evaluate the quality of the scanning. I also tested 10 words per volume to determine the efficacy of the text-search utilities offered by each of these resources. In addition, I examined the metadata for each volume and noted basic bibliographic data. I kept track of these findings for individual volumes on a spreadsheet.

The narrative portion of the report comprises a comparative analysis of the three resources over several criteria, and generally follows a sequential approach that reflects the order of steps in which the research was conducted: finding books; gathering data about them; viewing them; navigating through them; searching in them; and, finally, citing, repurposing, and sharing text. The report concludes with some thoughts about the strengths and weaknesses of these resources, problems to be rectified, and the promise and opportunities of mass digitization.

2. Scholarly Background/Methodological History

I received my Ph.D. in communication from the University of Pittsburgh, following a program in media studies that emphasized the history of mass media in the United States. I also earned a Ph.D. certificate in cultural studies. My dissertation, *The Humble Handmaid of Commerce: Chromolithographic Advertising and the Development of Consumer Culture, 1876–1900*, examines the social and cultural conditions under which the first visual advertisements for nationally marketed brand-name consumer products were produced and used. I received honorable mention in an award competition for best dissertation in media history from the American Journalism Historians Association in 2005. My published work includes a chapter in an anthology on the history of cigarette advertising in the nineteenth century (Stabile

2000). The early history of broadcasting, particularly radio, is also of interest to me although I have never written on the subject.

The primary-source research for my dissertation took me to archives and libraries in Pittsburgh, Washington, D.C., Cincinnati, and Boston. Much of the research is based on uncataloged lithography trade journals I found in the Printing and Graphic Arts division of the Smithsonian's National Museum of American History, where I was a predoctoral fellow for three months. Other primary source materials published in the late nineteenth century, such as lithographers' directories, advertising manuals, amateur art magazines and lithographed advertisements, were found in various divisions of the Smithsonian and the Library of Congress. Unpublished materials included lithographers' records (such as correspondence and workers' contracts) from historical societies in Pennsylvania and Ohio as well as at the Smithsonian. As this pattern suggests, as with any historical research, the methodological challenge is finding evidence in the form of records (broadly defined) of the historical events, activities, or social practices being investigated. Preferably these records will have been created contemporaneously with (or shortly after) the activities or practices in question by members of those social groups who are closest to the subject of inquiry.

2.1 Subject areas and search strategies

The three main disciplinary areas in which I conducted research for this project are:

1. British cultural studies (as identified with the "Birmingham School")
2. History of U.S. broadcasting, 1920–1950
3. History of consumer culture in the U.S., 1880–1920

An effort was made, for the purposes of this study, to expand relevant bibliographies of these disciplinary areas to include not only core texts and recent monographs but also a significant number of books available in full text. Since British cultural studies as a field is a contemporary one, dating its origins to research that was conducted in the mid-twentieth century at the University of Birmingham in Great Britain, an effort was made not only to find foundational texts in this field but also to identify some of the late-nineteenth-century antecedents that have been identified with it. These include works by authors such as nineteenth-century cultural critic Matthew Arnold, as well as by those thinkers who have provided many of the theoretical underpinnings for the field, notably Karl Marx and Frederick Engels. (I felt Marx's *Capital* to be an important enough work to justify examining three different digitized versions, and indeed I discovered a wide range of quality in the scans.)

Cultural studies is often concerned with the study of cultural forms associated with working class audiences. One text-based nineteenth-century cultural form that has been the subject of inquiry by cultural studies scholars is popular literature, in particular the dime novel (Denning 1987) or, as similar types of serialized

adventure stories were known in Great Britain, penny dreadfuls; I therefore searched for these types of books for this project. In addition, I searched for any kind of text that would provide insight into social attitudes toward dime novels, and I found advice, trade, and religious literature that provided perspectives from members of various professions.

For the history of broadcasting and the history of consumer culture, I began by searching for what I consider to be seminal texts in these subject areas, many of which were not found except for bare-bones bibliographic metadata on GBS. (I still included these in the bibliographies and indicated on the spreadsheet that they had evidently not been digitized and were not available in full or in part.) I then used citation searching and less structured surfing on GBS and MLSB to identify more books I would like to read. I found citation searching to be more useful for finding relevant scholarly books than keyword or subject searching (where available), and indeed I discovered many books of interest that were available in limited viewing format and would therefore be useful for this project. I ran out of time before I was able to examine all of the books I found. The full-text sources I found that are relevant to understanding the histories of radio and consumer culture included trade literature from the advertising and broadcasting industries, along with some consumer-oriented advice and political literature. I found these primarily by conducting keyword searches of terms such as *radio*, *broadcasting*, *consumer*, *advertising*, and *"pictorial advertising,"* as well as subject searches such as *consumer culture* and *radio broadcasting*, in GBS, MLSB, and IA. As described in section 3.1, not all of these options are available with all three of these resources.

3. Comparative Review of Three Resources

Following is an analysis of the three resources, focusing on their usefulness for scholarly and historical research. This section looks at how well search interfaces work to locate relevant materials; how much, and what kinds, of metadata about the works is provided; in what cases full text is provided and what kinds of limited access are offered for in-copyright works; how easy it is to get to various parts or pages of a digitized book; the types of OCR failures that occur preventing the successful location of search terms within individual texts; some difficulties that occur in trying to navigate around these sites; and utilities for sharing, repurposing, and citing text.

The hardware I used to conduct this research is nothing fancy: an Apple iBook 1.33 GHz PowerPC G4, with 512 MB built-in memory and a small 12-inch viewing screen. As for software, the operating system is Mac OS 10.4.11, and the Web browser is Firefox 2.0.0.13. The Internet access is broadband with a transfer speed of 1.5 Mbps for downloads.

3.1 Finding books

GBS and MLSB offer keyword searching that covers metadata and the full contents of books in these resources. Words or phrases (in quotes) can be searched. In IA, only metadata is searched.

3.1.1 Advanced searching

The option of an advanced search separates GBS from MLSB; the former allows the user to specify, for example, an author's name or the title of a book when searching. This is very useful, since, for example, searching an author's name returns results of books *about* that person in addition to books *by* them. If the person is prominent, there may be several books about them cluttering up the results when the object is to browse through titles by the author. For example, searching for *Raymond Williams* in MLSB returns in the first ten results nine books about Williams (or that mention him) and one book by him. It is very useful to be able to browse all available books by a certain author, particularly for people like me who sometimes have trouble remembering the exact titles of books. In GBS, the same search presents the user with the choice of finding books by Williams or that mention him. In GBS the advanced search feature can also be used to browse through titles by a certain author.

Among the other options available with the advanced search feature in GBS is a limit by date (range) of publication. This is very useful for historical research, and I found in GBS the results were usually accurate within the prescribed date range. Subject searching is also available, and these subjects seem to be taken from library catalogs. Although some of the subjects that are applied to books in GBS seem to be less than helpful, I was surprised to find that searching *dime novels* as a subject returned several books that cover this subject in depth. In MLSB, subjects can also be entered in the general search box. Searching the term *penny dreadfuls* resulted in hits of several books that matched this genre. Unfortunately, form/genre titles such as this are not usually applied in library cataloging as happened in this case. (A look at the relevant MARC records in UC-Berkeley's OPAC confirmed that this is where the term came from). It would be helpful for cultural studies research if form/genre subject terms were applied to more types of printed materials.

In IA, only metadata is searched. By using the general search box on the home page and restricting results to "texts" (or a more specific set of texts) it is possible to restrict results to digitized books. Although it is limiting not to be able to search full text, generally speaking, there is more complete metadata about each book to search within IA than with either GBS or MLSB. In some cases, there are user reviews attached to the metadata. This was useful for me in one case, because the book *In the Days of Washington; a Story of the American Revolution* (1896) was described in the review as, "Lots of action and hair's breadth escapes typical of dime novel material." This resulted in my finding this book by searching for key term *dime novel* even though this term was not in any of the "official" metadata. An advanced search option is also available in IA, with several options.

Among the most useful fields to search is subject, since there is no keyword searching of full text. I successfully used the subject field to find texts about consumer culture.

3.1.2 Ranking of results

In none of these resources was it evident to me exactly how resources are ranked. The ranking algorithms used by MLSB and GBS are no doubt both secret, but I noticed that in MLSB, the order in which results were presented was at times inconsistent. The same query run several times brought similar results in a slightly different order. For example, the query *women radio broadcasting advertising* (restricted to texts that are 100% viewable) brought these first four hits in this order:

Big Business and Radio (monograph,1939)
Radio Broadcast (periodical, 1925-26)
Outdoor, Street-car, and Radio Advertising (trade, 1936)
Radio Broadcast (periodical, 1926-27)

Run a minute later, the same query resulted in these top four hits:

A Decade of Radio Advertising (monograph, 1933)
Radio, the Fifth Estate (monograph, 1950)
Education by Radio (periodical, 1934)
The First Quarter-century of American Broadcasting (monograph, 1946)

In both cases, approximately all of the same hits appeared on the first page of hits, but in a slightly different order. The important point is that the exact order of the hits is not very significant. I ran this query in an attempt to find texts that covered radio advertising in the context of women (either as producers of advertising or as consumers). Each of the hits on the first page was clearly about radio. Searching each title individually was required in order to find out which ones actually discussed advertising and women in the same context. A significant number of hits on the first two pages did this, although they were not necessarily the highest hits. When their texts were searched individually, the titles earliest in the list of hits did all have the word *women* and the word *advertising* in significant numbers, but sometimes they were on different pages and therefore it was evident that advertising and women were being discussed in different contexts. It was therefore required to run the query *women advertising* for each volume to see which ones mentioned these terms on the same pages. This anecdote is not to point out a failure in the ranking algorithm, it is simply to suggest that the exact order of hits should not be taken too seriously, so whether the order of hits changes from one search to the next is perhaps not entirely significant.

3.1.3 Ease of using search results

GBS presents with each result some basic metadata and sometimes the beginning of a summary of the text (in the case of more recent publications) or a bit of text from the book including the search term

(in the case of full-view, public domain texts). In addition, the user is informed whether the book is offered in full view, a limited preview, snippet view, etc. With MLSB, the user is offered this basic metadata but can also conveniently mouse-over a result to find more information displayed to the right of the page. More metadata, in addition to a summary and/or table of contents, is often displayed for recent publications, and the user gets a visual display of how many times the search term appears in the book. IA results have basic metadata, but do not usually include the year of publication, which makes historical research very difficult (although limiting all results by date is possible with the advanced search feature and it is possible to sort results by date). It is helpful that sometimes the name authority for the author is displayed, which often shows his or her birth date. Sometimes a description of the physical characteristics of the digitized volume is offered in the results list, which is not very helpful. For example, the search result listing for *The village boys, or, Stories to persuade boys not to quarrel*, shows author and title plus: "Purple embossed cloth over boards, blocked in gold." While it is good practice to record this information, it is not usually what a user is looking for when browsing search results.

With MLSB, it can be difficult to navigate through a long list of search results because it only displays a select number of pages of results to click on. If you have 500 results on a search, you may want to skim through the first 50 results to get a feel for what is there. By the time you get to result 50, you're on the fifth page of results. If you then want to go back to the first page of results, you have to use the history on your browser, use the back browser button, or just run the search again because your choices are *prev*, *nex*t, and *pages 3-7*. This is not conducive to scholarly investigation; scholars must have the patience to browse many, many titles if necessary.

3.2 Metadata: descriptive and administrative

In doing scholarly research, it is vital to know, at minimum: the complete title of a book, who wrote and published the book, and the year it was published. In GBS, the user can view an "about this book" page—separate from the page where the book is viewed—which includes this type of descriptive metadata. For some historical works, I found this information was sometimes incomplete or even slightly inaccurate. For example, the record for Marx's *Capital* (1906) states: "Translated by Samuel Moore, Edward Bibbins Aveling, Ernest Untermann." However, the title page states that the translation was by Moore and Aveling; *the text was edited by Frederick Engels* and revised by Untermann. So unfortunately, the editor's name has somehow dropped out of the metadata. In addition, with both GBS and MLSB (on its "about" page, which is the same as its view-the-book page), it was disconcerting to find that, at least with my browser software, long titles are always cut off. So titles are displayed as *Marx's Critique of Political Economy: Intellectual Sources and* (GBS) or *Stuart Hall. Critical Dialogues in Cultural St* (MLBS). For long titles, then, it is necessary to examine the book cover image or find the title page to learn

the complete title to the book. In general, I found basic metadata such as the year of publication was sometimes missing in both GBS and MLSB.

On its "about" page, GBS often displays a summary of the book supplied by the publisher (which, as indicated earlier, MLSB offers by mousing over titles on the search results page). For out-of-copyright works, GBS sometimes offers "key words and phrases" in the book, which provides an at-a-glance way to determine if the book discusses the topic of interest. Clicking on a term activates the "search in this book" feature.

In IA, metadata is usually complete. However, when searching "texts," sometimes books digitized by Project Gutenberg are returned. Whether or not Project Gutenberg represents *mass* digitization, it is useful to have these results. Unfortunately, Project Gutenberg books don't always provide necessary metadata, even in the plain-text view (page views aren't available). For example, it required a search of GBS (although, admittedly, an OPAC would have been more reliable) to find out that the Project Gutenberg book *The Consumer Viewpoint* by Mildred Mattocks was written in 1920. On the other hand, this lack of metadata is the exception since most books returned by text searches in IA were digitized through the Open Content Alliance, and for these IA is very good at supplying metadata, including administrative metadata: it is always clear what library supplied the book, and often information about the condition of the scanned volume is provided. In contrast, with the MLSB and GBS, the name of the library is not always provided (it is often necessary to scroll through the book to find the stamp on the inside front or back cover) and no information about the book's condition is provided to the user.

3.3 Modes of access to the text

Generally speaking, for works in the public domain, MLSB and GBS allow for full-text viewing of contents. In IA, all books I found were available full text. For each of the commercial search services, in-copyright works are digitized through agreements with publishers. Publishers working with MLSB are given choices about how much of their books they want to be made available and the type of access provided to the user.

3.3.1 Full text

It is generally the case that digitized books published before 1923 are available full text and those published later (with some exceptions such as government publications) are not.

However, I found a large number of books and periodicals about radio published in the 1920s, 1930s, and 1940s available full text on MLSB. These were largely trade journals and other trade literature. It was perhaps even more interesting for me (and much less gratifying) to learn that not every book digitized by GBS that was published before 1923 is available in full text. For example, *The Art of Advertising: Its Theory and Practice Fully Described* (Stead, 1899) is provided only

in snippet format. I do not know if this is an error, or if there is some reason why Google felt this book was not in the public domain.

3.3.2 Limited view

My experience suggests that the most common way in which publishers have chosen to provide access to their books in MLSB is to allow users to view a certain percentage of the pages in each book (usually 10 percent, but at least one publisher, Oxford University Press, allows 20 percent). Users must create a Hotmail account and sign in every time they want to access a limited-view book, which I found worth the time. In my experience, there were no restrictions placed on which pages could be viewed, just the number that could be viewed. A convenient and useful service for users is MLSB's running page counter, which tells users how many pages they have left in their viewable allotment for that particular book.

GBS also allows publishers to limit the number of pages users can view, but in every limited-view book some pages seem to be designated as entirely off-limits. This could be just one single page in a chapter or hundreds of pages in a row, and this became a real problem when the endnotes to viewable pages were blocked or vice versa: when endnotes were available to pages that were blocked. While it is disconcerting to be blocked from certain pages, I found that the most problematic aspect of viewing a limited-view book in GBS was not knowing how many pages were left to view. At first, several chapters may be accessible (at least in part), but slowly I noticed more chapters becoming off-limits. (After several attempts I finally realized that the total number of pages allowed is usually 10 percent, as with MLSB.) I understand the object of the publisher program is not to accommodate a satisfying reading or research experience, but rather to provide the user with enough information so he or she will buy or borrow the book. However, I found it much easier to gather enough information about whether to obtain a book using MLSB's approach, because I could always read or skim the parts I needed to determine the book's usefulness, such as the table of contents, the introduction, the endnotes to whichever pages I was reading, the bibliography, and the index. To me, this is often just as important as being able to search for key words and terms throughout the text. For example, being able to glance through the entire introduction is important since scholarly monographs usually use it to summarize the main argument of the book and what each chapter covers.

Another way in which digitized books are made available on MLSB is by allowing "pages around search results" to be viewed. In this mode, the user can view the two pages immediately before and after the page that includes the term searched for, and there is still a limit to the overall number of pages that can be viewed. This is a fairly good way to provide access to some text in copyrighted works (at least random pages are not blocked). It is possible to gain access to the table of contents by searching for "contents," for example. However, it is difficult when it is hard to find the endnotes/references to the pages you are viewing.

3.3.3 Fragments and snippets

Another mode in MLSB is "page fragments viewable." Here again, the pages viewed are determined by search terms, but in this mode only part of the page is viewable. In the book *Broadcasting Freedom: Radio, War, and the Politics of Race, 1938–1948* (Savage 1999), I searched for the term *immigrants* and the results showed that 72 pages contained the term. Clicking on the first hit displayed a page with approximately three-quarters of it blurred out and the other quarter viewable. In the viewable portion, the term *immigrants* appeared three times, and elsewhere in the blurred-out portion of the page it was possible to see the term was highlighted another four times. I was therefore able to understand not only that the book discusses immigrants to a significant extent but also gleaned some of the context. However, without access to other descriptive data about this volume, the efficacy of this viewing mode would be limited for the purposes of determining whether to obtain the book. Fortunately, this particular book also had a summary and the table of contents available in the mouse-over view on the results page. Learning that this monograph contains a chapter titled "Americans all, immigrants all: cultural pluralism and Americannes" probably relates more about how the topic of immigration is approached than knowing the term is used in 72 different pages and seeing the context of some of these instances.

A final way that I experienced digitized texts being made available is with GBS's "snippet view." This contested mode is how copyright-protected books that have not been cleared for digitization by copyright holders are being made available to users. The intent is for GBS to present the user with three "snippets," of about four lines of text each, that all contain the queried term, in addition to providing information about how many total times the term is found in the book. I have found that in many cases, this viewing mode fails because the queried term is not even part of the snippet provided, the line with the term is cut off at the bottom of the snippet, or the snippet is even simply a blank section of a page or a header that does not include the term. For example, when I queried the term *labor* in Marx's *Economic and Philosophic Manuscripts of 1844* (1844, trans. 1964), two of the snippets showed page headers and the third showed the term used in a sentence from the introduction to the book. This last would have been mildly helpful, but it was disappointing that it came from the editor's introduction rather than the translation of Marx's writing. It is arguable whether being presented with snippets such as these is any more helpful than just knowing the number of times a term appears in a particular book.

3.4 Navigating the text/structural metadata

Although physical cardboard-and-paper books are sometimes considered to be conducive only to a linear way of reading a text, one of the reasons the codex has been such a successful technology is no doubt the way it allows the user to quickly page through it, moving easily from reading one part to another in a way not available with

previous formats such as scrolls. The usability of the digitized-book interface depends on respecting the nature of the codex, allowing the user to find and move to various sections throughout the book: the chapters, notes, tables, illustrations, index, etc.

In GBS, users can get access to page views, with page-forward and back arrows along with a scroll bar to navigate through the book. Various levels of magnification are available (helpful when the user's viewing screen is small), as is a full-screen viewing option. Full-text books can be paged through two pages at a time (open-book view), and users can go to a particular page number, which is very helpful. In addition to page views, viewing plain text is available with full-text books, allowing the user to see the text of each page as interpreted by OCR. Often, the table of contents is provided as a link on the "about this book" page, but in most cases there are very serious OCR problems in this metadata. In some cases, the table of contents in the page view contains clickable links to the relevant page numbers, but sometimes these links fail as well. The "read this book" page usually contains useful links to various elements of the book, including the table of contents, title page, and index.

The MLSB viewing experience is similar in some basic ways but different in many others. As with GBS, users can move through the page-view text one page at a time or can scroll through the book, and with MLSB they can use an enhanced scroll bar to easily move to pages that contain queried terms. For full-text books, MLSB usually has a link to the table of contents page, and often the table of contents includes links to the named chapters, although I found, as with GBS, there was usually at least one bad or missing link in every book. (In both resources, these link failures were either caused by OCR that failed to accurately read the page number or by the failure to apply a link at all.) MLSB has only two levels of magnification: fit-on-page and expand to fit the width of the viewing area; there is no double-page viewing option.

The most difficult aspect of using the MLSB interface is that the user cannot easily go to a particular page number, even with full-text books. This is important in many instances, but to give one example: often guides in the front matter, such as a list of tables, do not contain links to relevant pages. If the table a user want to see is on page 346, he or she cannot go directly to this page but must use the bar on the right to estimate the location of the page and scroll through from there. Another option is to try to find the table by typing its title in the "search inside the book" box. However, in cases when the table is arranged horizontally on the page (at a 90 degree angle so if you were holding the book you would rotate it to see the table), OCR often fails and the text may not be found.

In IA's flip-book viewing interface, which is the default of several viewing options, the user is presented with a graphical 3-D representation of the book. There is no magnification option (although this feature is evidently in the works), so the best option for preventing eyestrain when viewing a large amount of text with a small screen is the time-consuming process of downloading the .pdf. As the name

implies, it is possible with this interface to "flip through the book" forward and backward one page at a time. The user can also click on a page later in the "closed" part of the book to go there, although it is not possible to know precisely what page because the message that appears when the user mouses over a page indicates the number of the leaf they will be taken to, not the page as printed in the book. This can lead to confusion for the casual user who is not familiar with book-binding terminology, and is not particularly useful even for those who are.

3.5 Searching for text in books

The foregoing section discusses ways in which the various search interfaces make it possible for the user to employ the structural elements of the printed book, particularly the table of contents and page numbers, to navigate through the text of its digitized version. This section deals with how the interfaces maximize one of the capabilities made possible through digitization: key word searching.

MLSB, GBS, and IA all allow the user to search for terms in the text of the book while in page-view mode. In MLSB and GBS, the user is presented with the number of the page on which the queried term is found, along with a short fragment of the sentence or paragraph in which it appears. The user can scroll through and read these fragments to determine which page to click on. In MLSB, these search results are presented in order of some kind of relevance (number of times the term appears?), but the user can also use the enhanced scroll bar to scroll to pages that contain the term. In GBS, the user must scroll through terms in the order they appear in the book, with no option to see results in order of relevance. In both resources, queried terms are highlighted when the user clicks to the page.

In IA's flip-book mode, the pages that contain the queried term appear with a graphical representation of small sticky-note flagging pages that contain the term along with a very brief fragment of surrounding text. There is no relevance option, and often it seems the term is only highlighted one time per page, even if the term appears on a page more than once. Sometimes the search box becomes partially obscured by the sticky notes or the image of the book itself. It should be noted that another option for viewing digitized books on the IA is using the DJVU open-source reader format, which is a plug-in designed to work with IE, Firefox, or Safari browsers. Using DJVU, the user can page through the book and search for terms, but the searching feature is less robust and more glitchy than with the flip-book mode and I found it was not worth the time it took to download the plug-in. DJVU allows the user to page through the book going from one page to the next that includes the queried word, highlighting the word, but in at least one case I found it did not work properly. I queried a word that appears many times in the book (*gold* in *Real Gold: An Adventure Story*), and DJVU would find the next time it appeared but would not go on to the next and the next, and the search term box kept disappearing.

The ability to search for phrases rather than simply single terms

is vital. For example, if the subject of inquiry is the *National Association of Broadcasters* or *motion pictures*, the user cannot be expected to sort through all the pages in which each of these individual words appear in order to find the instances in which this phrase is presented. Both GBS and MLSB allow for phrases in quotes to be searched, but IA does not allow phrase searching in any of the viewing interfaces I tried, and the flip book sometimes even stalls for a complete minute or two when presented with a quoted phrase before giving a "not found" message.

3.5.1 Text quality/scanning/OCR problems

These are grouped together because they are related problems that can result in text-searching failures, and it is often not possible to determine the exact cause of these failures. Some of the detectable factors that lead to failures are related to OCR problems such as the lack of the application of OCR to particular pages or page fragments; marked-up text; and text that is at a 90 degree angle, such as in landscape-shaped tables and figures.

Sometimes OCR is not applied where it would be helpful, such as to the list of "principal characters" in the front matter to *Telecommunications, Mass Media and Democracy* (McChesney 1994) as digitized by GBS. This list contains frequently invoked names along with a brief biographical note about the person (e.g., Ernst, Morris: ACLU radio committee member). Lists like this can be helpful when reading complex histories of an industry. A similar OCR failure is the skipping of entire blocks of text (e.g., footnote #3 in *Capital,* 1906, p. 200).

In GBS, MLSB and IA, OCR is usually applied to text printed in the long/landscape format in tables, graphs, photo captions, etc. But often this OCR contains many errors. *Radio, the Fifth Estate* (Waller 1950) as digitized by MLSB includes 20 pages of NBC network programming schedules from the 1940s that contain vital information such as the name of the host, show titles, and advertisers/sponsors. OCR was applied to these schedules but only worked in a few cases. This could result in partial discovery of information: if the user is searching for shows sponsored by advertiser Prince Matchabelli, these could be found, but not shows produced by Kraft. If they wanted to know about the programs "Labor for Victory" or "The Catholic Hour," they would not find any information about them because the OCR failed repeatedly to pick them up. If they were interested in the host named Donovan, they could find out some of his or her schedule, but not most of it. (It would also be nice to be able to rotate the screen to read the schedules.)

Often, a word search failure can be diagnosed in GBS by switching to plain-text view and noticing that the OCR software has misinterpreted a word. A typical error would be for the word *dehumanization* to be read like *dehuuianization.* OCR fails very often when a text has heavy underlining or is otherwise marked up in such a way that the text is partially obscured. In my analysis, I checked 10 words/phrases per book to determine the efficacy of the scanning and OCR

for text searching. I spaced the terms checked throughout each book, but when possible I deliberately checked words that were both significant and seemed they would be difficult to pick up: small text, blurry-looking text, words in italics, words split at the end of a line of text, and text that had been marked up. My sense is that perhaps the most likely predictor of OCR failure is marked-up text. Particularly when text is sloppily underlined, this can interfere with finding terms. The irony is, of course, that the underlined passages are often important ones as judged by a previous reader! An example of OCR failing in the face of underlined text can be found on page 32 of the University of Michigan library's book *Advertising: Selling the Consumer* (Mahin 1914) as digitized by GBS. The underlined phrase, *He uses advertising in the same way*, is interpreted as the following as evidenced by Google's plain-text view:

> He uses.
> adverti'iinfT tn thp sgrpf* way"

However, many text-searching failures are not as easy to figure out. Even when OCR interprets the word correctly, a "search in this book" search is not always successful. For example, it is easy to see how OCR did not pick up *advertising* in the example offered above, but it is not clear why it did not offer in the search results the fact that *engraver* appeared on p. 220 or 230, when it did highlight the term and plain-text view showed it read the word correctly. This is just one example of many cases in which OCR seemed to succeed (often the search term is even highlighted), but the search for a term still failed to flag the page. Spot checks suggested this happened in all three resources.

3.6 Returning to search results

In most user-friendly scholarly research resources it is possible to easily navigate back to search results at any stage in the process. This is important because usually scholarly research involves reviewing or scanning a large number of search results to determine which ones are relevant and to make sure nothing important is missed. Therefore, the ability to easily navigate between page views, metadata, and search results is key.

In MLSB there is a link back to search results. In GBS, there is no link back to search results from either the "about this book" page or the "read this book" page. It is possible to use the browser's back button, but I found this less than satisfactory because in some cases I had been back and forth between the "about" and "read" pages a few times while examining the book. In the IA flip-book mode, it is not easy to get back to the metadata page or to search results. Using back browser, it is necessary to back through every page viewed and this can be time- (and patience) prohibitive. It is possible to click on the "details" button to get some metadata, but in order to get back to the search interface to conduct a new search it is necessary to use the browser's history. There is no link back to the search results from a metadata page.

3.7 Sharing, repurposing, citing

This subsection deals with how well these three resources allow for interactions with the text, sharing text, and keeping track of text and citations. Of the three resources, GBS has made the most attempts to move toward Web 2.0 functionality for its digitized books, but it is still a work in progress.

3.7.1 Bibliographic data management

GBS and MLSB both have links to "find in a library" from their "about" or "read" pages. This connects to the book's record in WorldCat, which has options for exporting the record to Refworks or Endnote. As a user of Refworks, I find this a very easy way to manage the bibliographic data for a large research project. It would perhaps be more user-friendly, however, if the link to WorldCat opened in a new window/tab, since it is sometimes necessary to do a bit of additional searching to locate the correct edition of the book (the user is not always immediately connected to the correct edition), after which the user may wish to return to GBS/MLSB and continue examining the book and searching for others. While virtually every book I found in GBS or MLSB had the "library" link, I noticed that one GBS book, *Radio Reader: Essays in the Cultural History of Radio* (Hilmes and Loviglio 2002), had links to booksellers but not to WorldCat. Upon further investigation I discovered that no books published by Routledge had the "library" link in GBS, so it appears to be an agreement with this publisher. However, I did find the "library" link in Routledge books digitized by MLSB.

3.7.2 Printing and copying

For books still in copyright, no printing option is available in GBS and MLSB. In these resources and IA, for full-text books in the public domain, it is possible to download a .pdf and print from it (one assumes, although I did not try this). It is inconvenient to be required to download the entire .pdf if the object is only to print out select pages.

I found no way of copying text in MLSB, but both GBS and IA have the option of viewing plain text in addition to page views, and this allows the user to copy and paste text into a new document. However, I found this to be easier in IA than GBS. I tried to use the plain-text option to copy an entire article ("Means of Leading Boys from the Dime Novel to Better Literature") from a digitized volume of *Library Journal* (Young 1896) but had trouble copying several complete pages of text because it was difficult to select more than one page at a time (the highlighting would disappear.) Selecting/cutting/pasting in chunks worked better, but it was easy to lose track of what had already been copied. It seems this technique would work fine for copying a paragraph or two to quote from the article, but not an entire article. Unfortunately, in order to retain a copy of this article, it would probably be necessary to download the entire lengthy volume of *Library Journal* (700+ pages) to obtain this one article.

3.7.3 Tracking and sharing

Perhaps the best way to deal with the problem just discussed would be to keep track of this article by "clipping out" its title and storing it in a Google notebook. This utility allows the user to capture stable URLs to specific pages in digitized books, give them descriptive labels, and return to them at any time. The notebook is designed so it can be shared with others. I did not test this sharing feature, but I did use the notebook to clip out a passage from a book and e-mail it to a friend. Another Google utility that is meant to facilitate sharing and collaboration is "my library," which I used at the beginning of my project but then stopped using. I did notice that "my library" not only allows sharing, but records can be exported in xml, which may make it possible to import them to bibliographic management software, although I did not try this.

3.8 Multivolume works and periodicals

One of the most vexing problems with the resources studied here is the failure to link separate volumes of multivolume works (or the failure to digitize all of the volumes of a particular work) and periodicals. This is frustrating for the user, who would benefit from links between different volumes that may be physically separate from one another—and so are treated as separate books for the purposes of mass digitization—but nevertheless share intellectual connections.

There are several problems with the digitization of multivolume works that it seems could be rectified with a minimum of attention to metadata clarity. In some cases, books that are one volume of a multivolume work do not include this information in the metadata. The *Encyclopedia of Radio* (2004) as digitized on MLSB actually consists of only the first volume, covering letters A–E. Since this book was digitized through the publisher program, it was evidently felt this was enough to provide the user with a feel for the book. Nevertheless, it would make sense for the volume number or coverage to be included in the metadata, which it was not. Other multivolume works have been digitized in their entirety, but the metadata makes it difficult to figure out that individual volumes make up a coherent whole. Erik Barnouw's *A History of Broadcasting in the United States* (1966) is a three-volume work or series, each volume having a different individual title. When searching in MLSB for the series title and the author's last name, the first three hits are these books, which is helpful. However, on the negative side, it could be confusing for the user who does not know it's a three-volume work, because the individual titles of the books are not displayed nor is there an indication of different volume numbers. Each title appears exactly the same in the results list; the initial thought is that these are different editions of the same text (unless you look closely at the tiny cover image to see the different titles). Ideally, links to the other volumes would be provided and individual book titles would be presented.

Dime novels were published serially, a problem that has led to some peculiarities and errors in their digitization. For example, in MLSB, a 22-page dime novel story titled *Pacific Pete, the Prince of the*

Revolver, which is evidently Vol. II, No. 18, of *Beadle's Dime Library,* is labeled with the series title *Beadle's Dime Library*. The story title is not part of the metadata record, nor is the author Jos. E. Badger Jr. While searching *Pacific Pete* does return the work in the results list, it could easily be missed because the user who wants to access this work may, understandably, not click on the title *Beadle's Dime Library,* assuming this is a book that only mentions Pacific Pete rather than being the full text of the work.

Periodicals are generally digitized as individual books, with no linking or relation between different volumes. GBS perhaps provides the best clarity in search results, which, at least in the case of *Library Journal*, indicate the year of the volume rather than simply the range of years the volume was published, as MLSB does. GBS's advanced search feature also allows a search to be limited to a particular periodical, and using the "other editions" link also accomplishes the retrieval of all volumes of the periodical. Ideally, the user would have the option of ordering results chronologically as well. MLSB provides in the results list and metadata the total range of years the periodical was published. So, for example, the Catholic periodical *Ave Maria* is given the date 1865–1970 in MLSB; the date of any individual volume must be ascertained by examining the text itself. There is no simple way to get a list of all volumes of this periodical and browse through them methodically, which a scholar would want to do if it was germane to his or her research. In IA every volume of *Radio Broadcast* is given the dates 1922–1930 in the metadata, but the individual volume numbers are also given and provided in the results list.

4. Analysis of Individual Volumes

An Excel spreadsheet, provided at the end of this summary, contains data about individual volumes. There are three worksheets, one for each of the resources studied: (1) Google Book Search, (2) Microsoft Live Search Books, and (3) Internet Archive.

5. Conclusion

Each of the three resources examined here has its individual strengths: GBS with its large corpus of works, its combination of full-text and advanced bibliographic searching, and the relative ease with which users can get around a full-text book; MLSB with its elegant-looking and functional results-list interface, transparency in informing users how much of a book they have left to read, ranked search results, and cleverly enhanced scroll bars; and IA with its serious attention to metadata (for its OCA projects) and its commitment to digitizing materials that can be made available full text. Each resource has its weaknesses as well: GBS's lower scanning quality (as suggested by some of the volumes examined for this study), choice of some highly marked-up texts, and failure to alert users when they will run out of pages left to view in a book; MLSB's lack

of an advanced search feature and its failure to allow navigation to particular pages of a book; and IA's failure to allow phrase searching and full-text searching across books in the collection.

The strengths mentioned here and the usefulness of mass digitization that I have only touched on in this report represent a very great benefit to scholars. It goes without saying that to be able to access books in faraway libraries and to be able to search contents across millions of books at far beyond the granularity made possible by print indices is a boon. I suspect this excitement is felt particularly keenly by those scholars who, like me, are interested in U.S. culture before 1923, since the amount of primary-source research material available is growing on a daily basis. And old methodologies can be adapted to this new medium, creating new opportunities. For example, many humanities scholars are fond of the serendipitous discovery made possible by shelf browsing. Mass digitization presents new opportunities for serendipity and new kinds of browsing. The designers of these resources could facilitate such methods of research even further in a number of ways, including linking together volumes of periodicals, thereby allowing researchers to browse through many volumes once they have, perhaps serendipitously, discovered the title. With many types of periodicals, like trade journals or other periodicals that give voice to a community, it is only by methodically reading (or at least browsing) many volumes that researchers are able to immerse themselves in that community, adopt their lingo, learn about their concerns. This is the type of laborious research that is not necessarily facilitated by full-text keyword searching, although it most definitely can be facilitated by bringing texts to the scholar's desktop through mass digitization.

So while there are great opportunities, there are also missed ones. The most disconcerting aspect of this research, in my view, is the failure to choose decent-quality, clean texts to digitize, and the related problems of low quality of scanning and resulting OCR. These problems, of all I have encountered, seem particularly acute because rectifying them must occur at the individual item level, reselecting and rescanning texts. Many of the other problems discussed here have their causes in inadequate interface design and algorithms, which can continually be improved. In other words, many of these problems can be solved at the level of the collection. The metadata problems discussed in this report seem to fall somewhere in-between these two: it may be that many problems—such as the lack of links between related volumes in periodicals or other multivolume works—do involve changing or enhancing data at the item level. Still, some of the most time-intensive aspects of mass digitization—selection, scanning, reshelving—would not have to be redone to make these enhancements to the metadata.

The main reason OCR failures and other text-searching failures are so problematic for mass-digitization projects is that users may have an overly confident view of how well these functions are working. It would be interesting to include in a study of users of mass digitization a question about what they believe the error rate is for

the books/terms they are searching, and to compare this to the actual error rate. My hypothesis would be that the users believe they are achieving a higher rate of success than they actually are. This is significant. One hopes scholars are not using the search-in-this-book function of mass-digitized books to do quantitative textual analysis; this would not be an accurate way to learn how many times Marx used the term *alienation* in his analyses of capitalist production.

It should be noted, however, that some of the texts I examined for this study are not representative of all digitized materials, and in fact may bring out the worst of it in some ways: I went looking for forms and genres of texts that are particularly ill-suited to mass digitization, notably dime novels. Like other cultural forms of interest to cultural studies in the Birmingham School tradition—those that are by, and/or for, working class audiences—they are highly visual, cheaply made, ill-preserved, and ephemeral. The challenges this creates for mass digitization can be met: images can be digitized and described, serial literature can be linked, and fragile paper can be partially stabilized through conservation work and treated with care. However, this would require more time to be spent per item than is generally the case with mass digitization and therefore may even strain the definition of *mass* digitization. Another factor in digitizing the products of low culture is the low quality of the printing, resulting in crooked pages, small margins, some blurred text, even typos/misspellings—illegibilities that carry over into the digital copy.

Searching for low-culture materials also serves to highlight the apparent ad hoc nature of some selections for mass digitization. *Pacific Pete* appears to be an example of this. The book bears the stamp of the UCLA library, and according to that institution's OPAC, it holds several volumes of *Beadle's Dime Library,* including 20 stories (volumes 1–3) of stories believed to have been published around 1876, that are evidently bound together. It is unclear, therefore, why only a single story was digitized, and this one is apparently not even intact—it ends on page 21 with what appears to be an incomplete sentence at the bottom of the page.

The sentence at the end of *Pacific Pete* reflects the action-packed nature of this dime novel, and is all the more suspenseful for being incomplete: "Springing forward, the hunter attacked the furious beast. …" Perhaps this is not too tortured a metaphor for the challenges that face the managers of mass-digitization projects as they attempt to bring under control the huge amount of untamed text that scholars are anxiously waiting to read. It is my hope that this report has illuminated some of the successes and failures of mass digitization, and how it could be improved to further the work of scholarly research.

6. Bibliography

Anonymous. 1927. What is the Matter With Radio Advertising? *Radio Broadcast* 12(1): 14-15.

Arnold, Matthew. 1869. *Culture and Anarchy: An Essay in Political and Social Criticism*. London: Smith, Elder and Co.

Arnold, Matthew, and William Savage Johnson. 1913. S*elections From the Prose Works of Matthew Arnold*. New York: Houghton Mifflin Company.

Arnold, Matthew, George William Erskine Russell, and Thomas Burnett Smart. 1903. *The Works of Matthew Arnold.* London: Macmillan.

Badger, Jos E. 1878. *Pacific Pete, the Prince of the Revolver*. New York: Beadle & Adams.

Barnouw, Erik. 1966. *A Tower in Babel: A History of Broadcasting in the United States, to 1933*. New York: Oxford University Press.

Beagle, Floyd L. 1920. *"Blood and Thunders" or, Dime Novels of the 80's and 90's*. Brooklyn, NY: C. Bragin.

Blanchard, Frank Le Roy. 1921. *The Essentials of Advertising*. New York: McGraw-Hill Book Company.

Boddy, William. 2004. *New Media and the Popular Imagination: Launching Radio, Television, and Digital Media in the United States*. Oxford: Oxford University Press.

Bottomore, T. B. 1991. *A Dictionary of Marxist Thought*. Cambridge, MA: Blackwell.

Cashmore, Ernest. 2004. *Encyclopedia of Race and Ethnic Studies*. New York: Routledge.

Columbia University. Bureau of Applied Social Research, Paul Felix Lazarsfeld, Harry Hubert Field, National Opinion Research Center, and National Association of Broadcasters. 1946. *The People Look at Radio*. Chapel Hill, NC: University of North Carolina Press.

Craig, Douglas B. 2000. *Fireside Politics: Radio and Political Culture in the United States, 1920–1940*. Baltimore, MD: Johns Hopkins University Press.

Cross, Gary S. 2000. *An All-Consuming Century: Why Commercialism Won in Modern America*. New York: Columbia University Press.

Curry, Charles Madison, and Erle Elsworth Clippinger. 1921. *Children's Literature: A Textbook of Sources for Teachers and Teacher-Training Classes*. New York: Rand, McNally & Co.

Denning, Michael. 1987. *Mechanic Accents: Dime Novels and Working-Class Culture in America*. London; New York: Verso.

Douglas, Susan J. 1987. *Inventing American Broadcasting, 1899–1922*. Baltimore, MD: Johns Hopkins University Press.

Egan, Maurice Francis. 1889. "A Raft in Mid-Ocean; or, Lessons Learned at Sea," chapter 11. *Ave Maria: Magazine Devoted to the Honor of the Blessed Virgin* 28(7) (Feb. 16): 165-168.

Engels, Friedrich, and Austin Lewis. 1903. *Feuerbach, the Roots of the Socialist Philosophy*. Chicago: C. H. Kerr & Co.

Fay, Lucy Ella, and Annie Thaxter Eaton. 1919. *Instruction in the Use of Books and Libraries: A Text-Book for Normal Schools and Colleges*. Boston: F. W. Faxon Co.

Fenn, George Manville. 1894. *Real Gold: A Story of Adventure*. London: W. & R. Chambers.

Fenn, George Manville. 1888. *Dick O' the Fens: A Tale of the Great East Swamp*. London: Blackie & Son.

Fowler, Nathaniel C. 1889. *About Advertising and Printing. A Concise, Practical, and Original Manual on the Art of Local Advertising.* Boston: L. Barta & Co.

Garvey, Ellen Gruber. 1996. *The Adman in the Parlor: Magazines and the Gendering of Consumer Culture, 1880s to 1910s*. New York: Oxford University Press.

Gramsci, Antonio, and Joseph A. Buttigieg. 1996. *Prison Notebooks*. New York: Columbia University Press.

Graydon, William Murray. 1900. *The Jungle Trappers: A Tale of the Indian Jungle*. London: J. F. Shaw.

———. 1896. *In the Days of Washington: A Story of the American Revolution*. Philadelphia: Penn Publishing Co.

Grossberg, Lawrence, and Janice Radway. 1996. *Controversies in Cultural Studies*. London: Routledge.

Hall, Stuart, et al. 1978. *Policing the Crisis: Mugging, the State, and Law and Order*. New York: Holmes & Meier.

Hall, Stuart, David Morley, and Kuan-Hsing Chen. 1996. *Stuart Hall: Critical Dialogues in Cultural Studies*. New York: Routledge.

Hall, Stuart. 1992. *Culture, Media, Language: Working Papers in Cultural Studies, 1972-79*. London and New York: Routledge in association with the Centre for Contemporary Cultural Studies, University of Birmingham.

Hall, Stuart, and Tony Jefferson. 1993. *Resistance Through Rituals: Youth Subcultures in Post-War Britain*. London: Routledge.

Halper, Donna L. 2001. *Invisible Stars: A Social History of Women in American Broadcasting*. Armonk, NY: M. E. Sharpe.

Herbert, Henry William, and James Jackson. 1882. *The Silent Rifleman: A Tale of the Texan Prairies*. London: James Jackson.

Hettinger, Herman S. 1933. *A Decade of Radio Advertising*. Chicago: University of Chicago Press.

Hilmes, Michele. 1997. *Radio Voices: American Broadcasting, 1922–1952*. Minneapolis, MN: University of Minnesota Press.

Hilmes, Michele, and Jason Loviglio. 2002. *Radio Reader: Essays in the Cultural History of Radio*. New York: Routledge.

Jackson, James. 1882. *Bottle-imp*. London: James Jackson.

Jackson, James. 1882. *Tom Terror, the Outlaw.* London: James Jackson.

Jacobson, Lisa. 2005. *Raising Consumers: Children and the American Mass Market in the Early Twentieth Century*. New York: Columbia University Press.

Jenks, George C. 1904. Dime Novel Makers. *The Bookman* 20 (October): 111-114.

Landry, Robert John. 1942. *Who, What, Why is Radio?* New York: G. W. Stewart, Inc.

Lazarsfeld, Paul Felix. 1971. *Radio and the Printed Page: An Introduction to the Study of Radio and its Role in the Communication of Ideas*. New York: Arno Press and *The New York Times*.

Loviglio, Jason. 2005. *Radio's Intimate Public: Network Broadcasting and Mass-Mediated Democracy*. Minneapolis, MN: University of Minnesota Press.

Lukacs, Georg. 1971. *History and Class Consciousness: Studies in Marxist Dialectics*. Cambridge, MA: MIT Press.

Mahin, John Lee. 1914. *Advertising: Selling the Consumer.* Garden City, NY: Doubleday, Page & Co. for the Associated Advertising Clubs of the World.

Martin, Mac. 1923. *Advertising Campaigns*. New York: Alexander Hamilton Institute.

Marx, Karl. 1964. *Economic and Philosophic Manuscripts of 1844.* New York: International Publishers.

Marx, Karl, and Friedrich Engels. 1900. *Manifesto of the Communist Party*. Chicago: C. H. Kerr.

Marx, Karl, Friedrich Engels, and Gareth Stedman Jones. 2002. *The Communist Manifesto*. New York: Penguin Books.

Marx, Karl, Samuel Moore, Edward Bibbins Aveling, Friedrich Engels, and Ernest Untermann. 1906. *Capital, A Critique of Political Economy.* New York: Modern Library.

Marx, Karl, and Nahum Isaac Stone. 1904. *A Contribution to the Critique of Political Economy*. Chicago: C. H. Kerr.

Marx, Karl, Friedrich Engels, C. P. Dutt, and V. Chattopadhyaya. *The Poverty of Philosophy.* With an Introduction by F. Engels. Moscow: Foreign Languages Publishing House.

McChesney, Robert Waterman. 1994. *Telecommunications, Mass Media, and Democracy: The Battle for the Control of U.S. Broadcasting,*

1928–1935. New York; Oxford: Oxford University Press.

McGovern, Charles. 2006. *Sold American: Consumption and Citizenship, 1890–1945*. Chapel Hill, NC: University of North Carolina Press.

McRobbie, Angela. 1994. *Postmodernism and Popular Culture*. New York: Routledge.

Mildred Maddocks. 1920. *The Consumer Viewpoint, Covering Vital Phases of Manufacturing and Selling Household Devices*. New York: Good Housekeeping Institute.

National Purity Congress, Aaron M. Powell, and American Purity Alliance. 1896. *The National Purity Congress, its papers, addresses, portraits. An illustrated record of the papers and addresses of the first National Purity Congress, held under the auspices of the American Purity Alliance ... Baltimore, October 14, 15 and 16, 1895.* New York: American Purity Alliance.

Peaslee, John Bradley. 1900. *Thoughts and Experiences in and out of School*. Cincinnati, OH: Printed for the author by Curts & Jennings.

Procter, James. 2004. *Stuart Hall*. New York: Routledge.

Radway, Janice A. 1984. *Reading the Romance: Women, Patriarchy, and Popular Literature*. Chapel Hill, NC: University of North Carolina Press.

Robinson, J. H., and James Jackson. 1882. *Pathaway, or Nick Whiffles on the Trail*. London: James Jackson.

Rutherford, Janice Williams. 2003. *Selling Mrs. Consumer: Christine Frederick & the Rise of Household Efficiency*. Athens, GA: University of Georgia Press.

Savage, Barbara Dianne. 1999. *Broadcasting Freedom: Radio, War, and the Politics of Race, 1938–1948*. Chapel Hill, NC: University of North Carolina Press.

Scott, W. D. 1912. *The Psychology of Advertising.* Boston: Small, Maynard and Co.

Shaw, Albert. 1901. The Place of Advertising in Modern Business. *The American Monthly Review of Reviews* 23: 612-614.

Shurick, Edward P. J. 1946. *The First Quarter-Century of American Broadcasting.* Kansas City, MO: Midland Publishing Co.

Stabile, Carol A. 2000. *Turning the Century: Essays in Media and Cultural Studies*. Boulder, CO: Westview Press.

Stead, William. 1899. *The Art of Advertising: Its Theory and Practice Fully Described.* London: T. B. Browne.

Sterling, Christopher H., Michael C. Keith, and Museum of Broadcast Communications. 2004. *The Museum of Broadcast Communications Encyclopedia of Radio*. New York: Fitzroy Dearborn.

Sterne, Jonathan. 2003. *The Audible Past: Cultural Origins of Sound*

Reproduction. Durham, NC: Duke University Press.

Strasser, Susan, Charles McGovern, and Matthias Judt. 1998. *Getting and Spending: European and American Consumer Societies in the Twentieth Century*. Cambridge, England: Cambridge University Press.

Sullivan, J. W. 1913. *Markets for the People: The Consumer's Part.* New York: Macmillan Co.

Thompson, E. P. 1966. *Making of the English Working Class.* [S.l.]: Peter Smith Publishers, Inc.

Tipper, Harry. 1919. *Advertising: Its Principles and Practice*. New York: Ronald Press.

Trent, William Peterfield. 1919. *A History of American Literature, 1607–1865.* Cambridge, England: Cambridge University Press.

Turner, Graeme. 2002. *British Cultural Studies: An Introduction*. London: Routledge.

Waller, Judith Cary. 1950. *Radio: The Fifth Estate.* Boston: Houghton Mifflin.

Williams, Raymond, and Ederyn Williams. 1990. *Television: Technology and Cultural Form*. London: Routledge.

Williams, Raymond. 2001. *The Long Revolution*. Orchard Park, NY: Broadview Press.

———. 1995. *The Sociology of Culture*. Chicago: University of Chicago Press.

———. 1983. *Culture and Society, 1780–1950*. New York: Columbia University Press.

———. 1976. *Keywords: A Vocabulary of Culture and Society*. New York: Oxford University Press.

Withers, Hartley. 1914. *Poverty and Waste*. London: Smith, Elder & Co.

Wurtzler, Steve J. 2007. *Electric Sounds: Technological Change and the Rise of Corporate Mass Media*. New York: Columbia University Press.

Young, Clement C. 1896. Means of Leading Boys from the Dime Novel to Better Literature. *Library Journal* 21(4): 144-147.