
A Critical Review

by

Thomas Mann

Prepared for AFSCME 2910
The Library of Congress Professional Guild
representing over 1,500 professional employees
www.guild2910.org

April 3, 2006

No copyright is claimed for this paper.
It may be freely reproduced, reprinted, and republished.

______________________________________________________________________________


The judgements made in this paper do not represent official views of the Library of Congress.

______________________________________________________________________________

Summary

According to the Calhoun report, library operations that are not digital, that do not result in resources that are remotely accessible, that involve professional human judgement or expertise, or that require conceptual categorization and standardization rather than relevance ranking of keywords, do not fit into its proposed “leadership” strategy. This strategy itself, however, is based on an inappropriate business model – and a misrepresentation of that business model to begin with. The Calhoun report draws unjustified conclusions about the digital age, inflates wishful thinking, fails to make critical distinctions, and disregards (as well as mischaracterizes) an alternative “niche” strategy for research libraries, to promote scholarship (rather than increase “market position”). Its recommendations to eliminate Library of Congress Subject Headings, and to use “fast turnaround” time as the “gold standard” in cataloging, are particularly unjustified, and would have serious negative consequences for the capacity of research libraries to promote scholarly research.
Karen Calhoun’s new report, The Changing Nature of the Catalog and its Integration with Other Discovery Tools, does read as though it were “commissioned by the Library of Congress,” as indeed it was. It articulates much of what has been the agenda of the current LC administration for several years now. A few snippets of its “Blueprint” recommendations indicate rather clearly its mind set regarding cataloging:

4.1.4 Manage acquisitions and catalog data through batch processes; as much as possible, avoid working on one record at a time

4.1.5 Identify local customization (e.g., for call numbers) and record editing practices and eliminate them in favor of accepting as much cataloging copy as possible without review or modification.

4.2.3 Abandon the attempt to do comprehensive subject analysis manually with LCSH in favor of subject keywords; urge LC to dismantle LCSH [This proposal is echoed in the pyramid presented on p. 14: “Simplify cataloging practice to a set of basic elements; eliminate LCSH]

4.3.5 Define fast turnaround and delivery of library materials to users as the gold standard of quality service, not the fullness of catalog data

[Italics added]

With that preview in mind, let us look at the Calhoun report in some detail.

The business model

It is instructive that the author regards research libraries within the overall framework of a business model:

“Each question [in the structured interviews] was intended to elicit the kind of information that an investor might want to know about any product or service (in this case, the catalog) whose market position is eroding. Ample evidence documents the declining market position of the library catalog,” (p. 8)
“The cost-effectiveness of cataloging tradition and practice is under fire.” (p. 9)

“The online library catalog has been a successful product. Like other products, it has passed through a life cycle.” (p. 10)

“... the existing local catalog’s market position has eroded to the point where there is real concern for its ability to weather the competition for information seekers’ attention.” (p. 10)

“The key problem of declining industries is the effect wrought by falling demand—that is, excess supply and capacity. . . . In healthy businesses, the capacity to produce a product matches the demand for the product.” (p. 11)

“Yet the key problem of today’s online catalog is the effect of declining demand. In healthy businesses, the demand for a product and the capacity to produce it are in balance. . . . research libraries investment in catalogs—and the collections they describe—does not reflect the shift in user demand.” (p. 15)

“Librarians will need a great deal more business acumen than in the past.” (p. 20)

“Stated in business terms, the library catalog can be said to be in a declining stage of the product life cycle ... Newer and more appealing products (like Google or Amazon) have entered the information market ....” (p. 26)

“The declining market position of the research library catalog puts research libraries on the horns of a dilemma. The problem is acute and cannot be ignored.” (p. 27)

There are two intertwined problems here. First, it is not appropriate, to begin with, to force research libraries into a business model because it does serious damage to their mission. And second, the business model that Calhoun proposes is itself unworkable as a business model.

As the above quotations indicate, Calhoun repeatedly uses the notions of “market position” and “supply and demand” as scourges for whipping traditional library catalogs. Apparently she sees increased market position within the information economy as the goal that research libraries ought to aim for. But this goal itself represents a substantive misunderstanding of a real business model.

In the real world (which I will define as that which exists outside hothouse library conferences) the goal of any business is to make a profit—which is not the same thing as the goal of increasing market share. That the two are different is illustrated by the fact that a business can still grow in absolute terms—i.e., regarding numbers of products shipped, number of employees hired, etc.—without increasing its market position if the overall market itself is growing.
Considerations of market position are a factor in business decision making, but they do not determine “the bottom line,” which is profit.

Research libraries are not in the “business” of earning financial profits in the market place. This is an elementary observation, but it seems not to have occurred to Ms. Calhoun. Indeed, the very funding that enables research libraries to continue in operation is not dependent on market place forces to begin with. Research institutions are traditionally funded by taxes levied by governments on citizens and by endowments; they are not funded by profits generated by their own operations. In general, they do not have to concern themselves with their “market position.” The whole purpose of relying on taxes and endowments, rather than self-generated profits, is to free research institutions from the marketplace forces of supply and demand.

Why does our society choose to fund its research institutions in a way that frees them from making profits—i.e., in a way that puts them outside a “business” model? I would suggest it is because the nation believes that they ought to serve a goal other than generating profits or occupying “market positions.” That goal is to be found in an academic rather than a business model: research libraries are funded to promote scholarship. And they are funded to promote it whether or not it turns a profit according to supply and demand criteria.

In the academic model there is an important concept called “tenure” that is lacking in the business model. Tenure is a mechanism for assuring that responsible scholarship is accomplished even if its objects and results would not be supported in the business model, because the expense of pursuing the research would not be justified by considerations of “demand”—i.e., because the results of the inquiry would not meet with enough “buyers” to support it. For example, for many years I, as a reference librarian, have been helping an academic who is compiling an exhaustive annotated bibliography on a particular Russian composer, Aleksandr Scriabin—whose name would probably not even be recognized by the majority of Americans. No advance from a commercial publisher would pay for the decades of time and sustained effort that the project has already required; sources from libraries and archives on three continents are included, and the more than 2200 citations are in multiple languages: English, Russian, French, German, Spanish, Italian, Romanian, Czech, Polish, Ukranian, Swedish, Finnish, Lithuanian, Japanese, Bulgarian, Turkish, and Portuguese. (An additional challenge: the composer’s surname can be transcribed or transliterated at least six different ways in the roman alphabet alone; his given name also has multiple variations. [Question: would a keyword search using Google’s software round up such material?]) The work continues because of the assurance of her tenure, and because her institution does not require her scholarship to be judged by market place standards.

Research libraries can be said to be institutions that have tenure. I have observed many thousands of scholars at work over nearly thirty years. If the research libraries that provide the obscure, the out-of-the-way, and the little-demanded material—as well as the vast stores of copyrighted books that are not digitized, and that are available only onsite—if these research libraries were held accountable, according to a business model, for turning a profit or for
increasing their “market position” in relation to the Internet, then the materials of scholarship, and the methods of gaining systematic access to them, would vanish in short order. That is why research libraries are supported so heavily by mechanisms that place them outside market forces of supply and demand. Our nation has the basic faith that scholarship in all areas ought to be pursued, whether or not it turns a profit—and in spite of the fact that the scholars exploiting the research libraries form only a very small percentage of the population.

Ms. Calhoun’s initial mistake—one which colors everything that follows in her report—is that of imposing a business model (whose goal is actually profit, not market position) on research libraries, rather than viewing them within an academic model (whose goal is the promotion of scholarship)—one that has been endorsed by centuries of practice within this country—which purposely positions such libraries outside marketplace demands.

**Introducing a misleading concealed proposition**

Calhoun’s use of the business model shows yet another distortion. It is that of misusing a metaphor: “The online library catalog has been a successful product. Like other products, it has passed through a life cycle” (p. 10). Notice the clever introduction of a concealed proposition: if the product is tied to a “life cycle,” the biological metaphor silently slips in the notion that death is inevitable. Elsewhere, Calhoun magnanimously admits that the dying catalog—whose “market position” is “eroding,” and whose fundamental value has been “challenge[d]” by “library leaders at LC and elsewhere”—“seems likely to continue for at least a couple decades and probably longer” (p. 5). But her generous concession, here, of a bit more future life in the catalog is cancelled on page 10: “Fortunately,” she says, “there are ways to use the knowledge that today’s catalog has reached the end of its life cycle” (p. 10). Excuse me? When did wishful thinking and rhetorical distortion of metaphors become objective “knowledge”? A dubious biological metaphor with a concealed proposition of finitude has now suddenly transformed itself into a conclusion that the “life” of today’s catalog “has reached”—not, apparently, a couple decades in the future, but already—its “end.”

The evident inability of the catalog to die quickly, however, is puzzling to Ms. Calhoun, because it persists, she points out, “notwithstanding the widespread expansion of digitization projects, ubiquitous e-journals, and a market that seems poised to move to e-books” (p. 5). How can such a dinosaur lumber on without realizing that its heart has already stopped beating—killed by digitization projects? (Calhoun’s remark here is remarkably similar to one made by Deanna Marcum, Associate Librarian of Congress; according to the distributed minutes of LC’s March 24, 2004, Cataloging Management Team Meeting, “[Marcum] said that Google was seeking agreements with libraries to digitize the content they owned. If Google succeeds in digitizing a library of eight to ten million volumes and making the content available on the Web, what happens to every other library? She said Library of Congress cataloging would not be needed in these circumstances” [italics added].)

The belief that “digitization projects” somehow kill the need for catalogs, or at least
hasten the end of their “life cycle,” is yet another unargued and undemonstrated proposition. The implication is that just throwing more keywords into the hopper eliminates the need for controlling search terms through standardization and authority work. And the further unargued implication, on top of that, is that if there are enough words available, then anyone can find something on any topic—and “something” is all that anyone needs.

Has “today’s catalog” reached its “end” just because there are now more keywords that can be searched? And are the needs of scholars essentially the same as the wants of quick information seekers? And does a “market” that “seems” poised to move to e-books really provide a death blow to traditional libraries?

The e-book market

Let’s return for a moment to “the market”: Is even the business world, let alone the academic, in fact moving in the direction asserted by Calhoun? Have people truly overcome their reluctance to buy or read e-books? Barnes & Noble stopped selling them in September of 2003 because there was no “demand” for them in the “market” whose verdict Calhoun elsewhere regards so highly. And a recent (2/21/06) article in the Washington Post’s Express edition, “Trying to Crack the E-Book,” reports the following:

GOING DIGITAL / Brown University junior Stuart Thompson jumped at the chance to save $30 and become a digital pioneer when his school offered a discounted, electronic version of a history textbook.

But after making the purchase, he noticed a few things amiss: He couldn’t run a highlighter over key points or jot notes in the margins, nor could he curl up with the tome without printing out the pages.

So much for the belief that this generation of youths wants everything digital: The publishing industry has been talking about electronic textbooks for a decade already, but sales remain minuscule.

Of the 100 or so students enrolled in the course where Houghton Mifflin Co.’s “A People and a Nation” is assigned, Thompson was the only student to buy the electronic version.

Brown is offering two other e-textbooks this semester, but none of the 150 students in those classes has bought one, even at 35 percent off the price of a new printed copy.

One might reasonably conclude that even the business “market” is in fact quite different from Calhoun’s depiction of it. E-books continue to meet with stiff resistance. (Perhaps Calhoun could allay our persistent belief in their unreadability by mentioning a half-dozen or so e-books that she has read herself. If she is reading them routinely, then undoubtedly the rest of us are behind the curve. But if she is not, then perhaps she might be a bit more reluctant to impose on
everyone else an expectation to “evolve” towards accepting a format that she herself finds uncongenial?)

Disregarding, and misrepresenting, the niche model

Apart from choosing a wrong model that entails a wrong goal for research libraries, and apart from also misrepresenting the actual goals of that wrong model, and apart from introducing a false concealed proposition of inevitable death in a “life cycle,” Calhoun’s report further distorts the options available to library administrators. She offers four that are derived from “Harrigan and Porter”—whose article, listed in the bibliography, is entitled “End-game strategies for declining industries.” (Do I detect a pattern of “agenda’d” perception here?) The four “strategic alternatives” that library administrators can follow, she says, are “leadership, niche, harvest, or quick divestment.” We need not concern ourselves, any more than she does, with “harvest” or “quick divestment.”

But the “niche” option is quite sensible. Calhoun, however, dismisses it in a single sentence—obviously her desired strategy is “leadership.” And she also misrepresents the niche option as much as she misrepresented the business model: “The niche strategy would be characterized by specialization; the research library or service organization would choose one or more segments of users—say, humanists or area studies specialists—in which demand is expected to be reasonably stable and move to serve those segments exclusively” (p. 12). This is a severe distortion of the option. The examples of “segments of users” given here, “humanists or area studies specialists,” makes it seem that the users to be reached by a niche strategy are to be segmented by subject areas—and, on top of that misinformation, that research libraries would then “move to serve those segments exclusively,” as though the libraries would thereby retreat from any concern for serving a larger population with general interests that lie outside the niche of those subject-defined “segments” of users.

The actual niche audience to be served by research libraries, however, is that of scholars—in all subject areas and disciplines—as distinct from quick information seekers. The crucial distinction that Calhoun misses is between levels of research rather than subjects of research. There is of course overlap between the scholars and “information seekers”—and, indeed, the very same people can have scholarly goals in mind for one project whereas they might need only quick information on another. But research libraries have a particular responsibility to serve the needs of scholars—especially because the alternatives of Google, Amazon, AltaVista, while being excellent sources for “something” on a subject provided “quickly,” nonetheless fail to support scholarship in several specific (and very important) ways.

The differences between scholarship and quick information seeking

Scholars, if I may be so bold as to draw on my own experience of working with thousands of them over a very long period, have these concerns that are not shared by quick “information
seekers”:

1) Scholars seek as clear and extensive an overview of all relevant sources as they can achieve.

2) They are especially concerned that they do not overlook sources that are unusually important, significant, or standard in their field of inquiry.

3) They do not wish to unnecessarily duplicate prior research.

4) They particularly wish to be aware of cross-disciplinary connections to their work.

5) They wish to find current books on a subject categorized with the prior books on the same subject, so that the newer works can be perceived in the context of the existing literature—not just in connection with the much smaller subset of titles that are currently in print. (Advanced scholars also wish for similar categorization of English language books with relevant foreign language titles, so that a worldwide context of literature on their subject can be easily discerned.)

6) They particularly appreciate mechanisms that enable them to recognize highly relevant sources whose keywords they cannot think up in advance, to enter into Boolean combinations in a blank search box.

7) Although they are more cognizant of the need for extended effort and the need to check multiple sources beyond the “first screen” of any Internet retrievals, they also wish to avoid having to sort through huge lists or displays—from any source—in which relevant materials are buried within inadequately-sorted mountains of chaff having the “right” keywords in the wrong conceptual contexts.

Google-type “relevance ranked” keyword searching cannot solve these problems; in fact, it exacerbates all of them. It is not even particularly good for keyword searching, in comparison with other databases—i.e., it does not allow full Boolean combinations in nested parentheses, or wildcard truncation, or proximity operators. Nor does Google-type searching allow searchers to limit the appearance of words to particular fields, because it does not distinguish or segregate such fields (title, contents, notes, bibliographies, etc.) to begin with. It just jumbles everything together.

Although she does sometimes use the word “scholars,” Calhoun does not use it in any way that clearly distinguishes its referent from quick “information seekers.” And her comments about researchers in general, while not untrue, are nevertheless slanted in a certain direction. Here are a sampling of her statements:
“... the existing local catalog’s market position has eroded to the point where there is real concern for its ability to weather the competition for information seekers’ attention.” (p. 10)

“At the same time, a large and growing number of students and scholars routinely bypass library catalogs in favor of other discovery tools, and the catalog represents a shrinking proportion of the scholarly information universe.” (p. 5, repeated on p. 9)

“As information seekers increasingly turn to search engines, research library leaders need to examine ways to bring the capacity to produce local online catalogs back into line with demand for them.” (p. 11)

[One of the] “challenges to feasibility” [is] “inability to base priorities on how users behave and what they want.” (p. 13)

“... search engines are students and scholars’ favorite place to begin a search. More users bypass catalogs for search engines, but research libraries’ investment in catalogs—and in the collections they describe—does not reflect the shift in user demand.” (p. 15)

“These developments have lowered the opportunity costs of obtaining information, increased self-sufficiency, and kicked off an accelerating decline in information seeking methods based on printed sources.” (p. 23)

“Speaking of the limited scope of the catalog and its emphasis on print, Norm Madeiros wrote ‘more and more, users want, expect, and pursue full text. In increasing numbers they look past the catalog when searching for e-journals, databases, and Web sites.’” (p. 25)

“Similarly, OCLC research on perceptions of libraries and information resources confirmed ‘the library is not the first or only stop for many information seekers. Search engines are the favorite place to begin a search ....’” (p. 26)

“... a large and growing proportion of students and scholars are bypassing the catalog in favor of other information resources, and the catalog represents a shrinking percentage of the scholarly information universe.” (p. 27)

“The most recent OCLC report presents compelling evidence that college students begin looking for information on search engines—89% of this group said they begin searches with a search engine vs. 2% who start their searches on library Web pages.” (pp. 36-37)
“An information seeker’s first exploratory point is highly likely to be outside the catalog.” (p. 37)

The “slant” here is on how frequently “information seekers” “bypass” library catalogs—or libraries themselves. In many cases, admittedly, Calhoun includes the qualification that users don’t use catalogs as their first exploratory point—in others, though, the message seems to be that users skip catalogs (or libraries) entirely. The conspicuous omission here is any of the abundant information pointing out how many “students and scholars” continue to use their local libraries regularly even though library use is not their first tack. A survey of these studies is included as an Appendix to this review. I suspect anyone who skims it will come away with an impression that researchers are not nearly as averse to using traditional libraries as Calhoun would wish us to believe.

What user studies actually say

Perhaps the gist of all of the user studies is summed up by the new Perceptions of Libraries and Information Resources from OCLC (2005) available at <http://www.oclc.org/reports/2005perceptions.htm>. Although its focus is more on public than academic libraries, it is still of interest. Let me quote excerpts from a review of this study, entitled “Sanity Check,” by Dick Kaser in Information Today (February, 2006), p. 16:

People are always extolling the virtues of libraries going virtual ... getting digitized ... going beyond the four walls. But you know what? In a recent market study by OCLC, Internet users seem to be very happy going out to the library and getting some good books. . . .

The study also indicates that most people think of a library as a physical place. In fact, they think of it as a place to go for books—not “information access,” not databases or other electronic resources—but to look at and check out good old-fashioned hardcover books printed on paper.

According to the study, when people have a specific question, they go to Google first for information. . . .

So what does all of this mean? OCLC’s researchers were troubled that “library” (as a brand) still means “books” to most people. They suggest the importance of using more outreach, education, and “product” positioning to change people’s outlooks. But that attitude presumes that books do not have—and will not have—enough enduring value to sustain libraries as institutions for readers.
From the survey results, nothing would seem further from the truth. Books are holding their own, and libraries could do worse than be perceived as the place to go for them.

Of all the industry players, book publishers should feel most heartened by this report. People are seeing books as having enduring value in the digital age. Indeed, the biggest complaint of those surveyed was that library books are not always current or in good condition. But the bottom line—and the great news—is that people still want books to read when they’re not searching Google.

**Scholarship as the proper niche for research libraries also serves quick information seeking**

Although it is the responsibility of research libraries to promote scholarship—as opposed to both “market position” and “quick information seeking”—the “niche” defined by scholarship, again, is not to be segmented “exclusively” to any subject or disciplinary slice of the pie. Research libraries have the professional responsibility to provide the capacities to scholars that will result in the deepest levels of research, in any subject area.

To focus on scholars as the “segment” to be aimed at, however, does not entail the disregard of “information seekers” who want only more superficial resources. The same resources need to be made available to all, even though they will not be used by all, or used to the same depth or purpose. In other words, a concern for the niche of promoting scholarship does not in the least exclude service to quick information seekers. The latter will simply find less than the former. They will be less systematic and more desultory.

For example, LC collects thousands of specialized subject encyclopedias whose texts are not online. The same sets can serve as both the end-point of superficial research or the starting point of more extensive scholarly research—especially when they are used in preference to the unreliable online Wikipedia. The same resources serve both user groups, but at different levels.

More to the point, the online catalog itself functions in a similar manner: an information seeker who wants fast access to “some” books on the “Bay of Pigs” invasion can do a keyword search on that phrase in LC’s OPAC and quickly come up with a list of 51 sources in English. (The results will actually be a mixture of books, sound recordings, and photographs.) A scholar of the subject however, can use the LCSH term (Cuba–History–Invasion, 1961) to achieve much more comprehensive results (86 entries, dozens of which do not use the keywords “Bay of Pigs”)—and in all languages (including Spanish, German, and Russian).

**The superiority of LCSH, over relevance ranked keywords, for scholarship**

Moreover, the structure of presentation provided by LCSH subject strings is much more intelligible than that provided by “relevance ranked” keyword displays—a point Calhoun seems not to notice. Specifically, a researcher using the Subject Browse (rather than Keyword) search
option in LC’s catalog will see following menu display; preceding each line is the number of records available under that aspect of the topic. Such a conceptual map enables the viewer to get not just a sense of the “shape” of the subject, but also of the “depths” of its different parts:

| 44 | Cuba History Invasion, 1961 |
| 4  | Cuba History Invasion, 1961 Aerial operations |
| 1  | Cuba History Invasion, 1961 Anniversaries, etc. |
| 1  | Cuba History Invasion, 1961 Congresses |
| 3  | Cuba History Invasion, 1961 Drama |
| 5  | Cuba History Invasion, 1961 Fiction |
| 6  | Cuba History Invasion, 1961 [from old catalog] |
| 1  | Cuba History Invasion, 1961 Juvenile literature |
| 14 | Cuba History Invasion, 1961 Personal narratives |
| 1  | Cuba History Invasion, 1961 Personal narratives, American |
| 2  | Cuba History Invasion, 1961 Pictorial works |
| 1  | Cuba History Invasion, 1961 Poetry |
| 1  | Cuba History Invasion, 1961 Prisoners and prisons |
| 1  | Cuba History Invasion, 1961 Regimental histories |
| 1  | Cuba History Invasion, 1961 Songs and music |
| 3  | Cuba History Invasion, 1961 Sources |

_Scholars_ who are moving into unfamiliar subject areas very much want to first get this kind of _overview of the shape and depth_ of the book literature within it. LCSH browse displays like this one provide exactly that kind of overview—one that is both _comprehensive_ and _comprehensible_. An array like this gives researchers both a reasonable outline of the _boundaries_ of the topic (within book literature) and a map of its _component parts_—neither of which is possible with either “relevance ranked” keyword searching or Amazon-type displays of “people who liked this also like these.”

Keyword searching of titles and tables of contents, no matter how the results are ranked, cannot possibly provide such an overview because it cannot retrieve, to begin with, all of the relevant English language records, let alone the scores of non-English books. Records that are not found cannot be ranked.

It is especially distressing that the Calhoun report is utterly oblivious to _scholarly_ work that _needs to incorporate foreign language books_. “Information seekers” in the Calhoun world are apparently assumed to want only English materials—which begs the question of why research libraries spend millions of dollars to collect foreign works in the first place, or bother to make available to U.S. scholars the works of their colleagues throughout the rest of the world. Under Calhoun’s proposals, such resources would drop out of sight to keyword searches because the
LCSH vocabulary control that collocates both English and foreign works under uniform subject headings would not be there to begin with. (One is reminded of what the angel tells George Bailey in *It’s a Wonderful Life*: the men on the troop ship drowned because George’s brother wasn’t there to save them; and the brother wasn’t there because George himself was missing. In these days of globalization and interpenetrating national concerns, do research libraries really wish to write off efficient access to foreign language books, or cause them to be noticed only with great difficulty by scholarly researchers? If so, welcome to academic Pottersville.)

Amazon-type searching, too, can give only “something” of the subject literature—i.e., scholars cannot rely on previous searchers’ whims to map out the *full extent of the aspects* of a subject in any systematic way because prior searchers who are heavily dependent on keyword searching, to begin with, cannot themselves see that range. It takes professional-level deliberation, intelligence, expertise, and craft to map out such ranges—and it is dangerously naive to think that the professional work of subject catalogers can be matched, or even approximated, by an “invisible hand” somehow guiding the random choices of multiple users at varying levels of interest, search skill, and knowledge of foreign languages. When the majority of relevant sources are not found due to keyword and “more like this” searches, the result is research that is haphazard, incomplete, and desultory. While that may be entirely adequate in serving the wants of “information seekers” who want only “something” provided “quickly,” it would directly undermine the capacities of substantive scholarship.

One might use a cooking analogy. Many people are probably quite happy with whatever cooking they can do, quickly and easily, on a George Foreman Grill. But others will need the additional capacities of stove-top burners and ovens for more complicated meals. Should *all* cooking be reduced to George Foreman grilling because it provides “something” in a way that is *faster* than an oven can produce? The point for libraries, and library catalogs, is that dumbing down the very capacities of the retrieval system for one segment of the population—“information seekers”—also dumbs it down for everyone else, including the scholars who need the browse displays such as that exemplified above, not just “relevance-ranked” records for English language books.

**The further superiority of vocabulary control over keywords for scholarly retrieval**

Note that the *control* of relevant literature brought about by LCSH is accomplished at two different levels—and Calhoun shows not the slightest awareness of either. First, the *vocabulary control* of the headings, to begin with, rounds up all of the works on the subject no matter what keywords they use. The 15 books under the **Personal narratives** and **Personal narratives, American** subdivisions, for example, include these titles:

- *Relatos de Giron*
- *Giron en la memoria*
- *Operacion Puma: la batalla aerea de Bahia de Cochinos*
- *Operation Puma: the air battle of the Bay of Pigs*
None of these works would be retrieved by a Boolean combination of the English-language keywords “Bay of Pigs” with either “first-person accounts” or “eyewitness accounts” or “primary sources” (which is how students usually ask for such things). And yet the collocation of works brought about by entering them all under a standardized heading enables researchers to see a whole variety of directly-relevant individual titles whose keywords they did not ask for. A quick-information seeker would never be able to specify in advance the range of words, from either titles or tables of contents, that would retrieve this set of records. Vocabulary control is necessary to create the conceptual category (Personal narratives) of these records—mere “relevance ranking” of their keywords would scatter these related titles to the winds.

Second, the browse display of left-anchored LCSH terms brings aspects of the subject to the attention of researchers that they would not think, beforehand, even to exist. How many people would think of options like Prisoners and prisons or Drama or Poetry in connection with “Bay of Pigs”? Who would think of typing those terms into a blank search box? The virtue of the LCSH system, and the browse displays of left-anchored subject strings that it creates in OPACs, is that it very easily enables researchers to recognize not just relevant individual titles that could not be specified in advance, but whole conceptual categories that are equally unanticipated, but nonetheless relevant to anyone trying to get an overview of an unfamiliar topic.

In writing The Oxford Guide to Library Research I found it particularly necessary to concentrate on the methods of searching that enable scholars to get overviews of the literature of their subject—because that is where their greatest problem lies. The more that information resources increase exponentially, the harder it is for anyone to see “the shape of the elephant.” Even without the presence of the Internet, researchers are still routinely overwhelmed even by OPAC searches if they use only keywords rather than LCSH mechanisms.
It is noteworthy that a major study of subject cataloging explicitly recommends the maintenance of left-anchored browse displays of LC subject strings. It is “Recommendations for Providing Access to, Display of Navigation within and among, and Modifications of Existing Practice Regarding Subject Reference Structures in Automated Systems,” (LRTS 49 [2005], 154-66), from The Association for Library Collections and Technical Services (ALCTS) Cataloging and Classification Section Subject Analysis Committee (SAC). This study is the product of nearly ten years’ work by three SAC subcommittees, charged (in its own words) with “investigating the theoretical, pragmatic, and political dimensions of improving subject access through better use of reference structure data.” Calhoun does not even mention this work.

Similarly, she avoids the papers of the Bicentennial Conference on Bibliographic Control for the New Millennium that would challenge her agenda, by the simple expedient of asserting, “Since the bicentennial conference, library leaders at LC and elsewhere have pushed beyond the questions posed at that conference to challenge the value of the catalog itself.” Pushed beyond—?? A more reasonable interpretation is that the “library leaders at LC and elsewhere” have simply not bothered to confront papers that challenge their agenda with inconvenient facts, arguments, and examples.

Legitimate problems with LCSH - the need for prior instruction to do any effective research

Am I asserting that there are no problems with LCSH? Of course not. It may be objected that, for example, no user would think of the LCSH term Personal narratives rather than the phrase “primary sources”—and this is indeed a valid point. But it is also easily answered. The fact is, no researchers—either scholarly or superficial—will ever do efficient searches in online resources without some prior instruction and education. (After a quarter century of standing over thousands of their shoulders at computer terminals, I have found remarkably few exceptions.) Left to their own devices—i.e., without any prior instruction or education—they will always find only “something” rather than an overview of the full range of material available to them. This, in my experience, is just a “given” that has more to do with human nature and the Principle of Least Effort than it has to do with the complexities of LCSH—or of the different but equally real complexities of keyword searching, citation searching, related record searching, efficient browsing, or any other search technique. “Under the hood” or “behind the scenes” programming, no matter how clever it may be, will not round up “everything” or make these complexities and trade-offs disappear. Keyword searching, just because it is most people’s “default” search technique, cannot realistically be considered “simple,” especially if the goal of the host research library is to solve the seven problems of scholars listed above.

One question, then, is how best to provide instruction—even minimally. And the fact is, it is much easier to teach browse displays of LCSH subject strings, and do it quickly, than it is to teach “critical thinking in evaluating Web sites.” The LCSH strings do not have to be specified in advance, beyond the initial term in the string—in a properly constructed OPAC they all show up...
automatically to provide the structured roadmap of options within a topic. Once they’re on the screen, the users can generally proceed very efficiently from there—it’s a matter of just recognizing what’s in front of them, from a structured menu. “Critical thinking,” in contrast, is not something that can be taught quickly or easily under any circumstances—learning to do it requires a great deal of writing, and correction of that writing, to take hold; and it is usually not the job of librarians in research libraries to correct term papers, even those papers involving library use. As to the effectiveness of Google-type automatic “relevance ranking,” I will leave readers to their own experience to judge how adequately and efficiently it sorts for scholarly overview purposes the tens of thousands of records retrieved in full text databases (or even in OPACs).

“Fast turnaround” of cataloging over quality

It will be noticed that I did not list “speed” of access as one of the seven critical problems that scholars face; and in doing so I am directly challenging Calhoun’s recommendation “4.3.5 Define fast turnaround and delivery of library materials to users as the gold standard of quality service, not the fullness of catalog data.” (I cannot help but detect here the influence of Calhoun’s colleague at Cornell, Sarah Thomas, the former Director of Cataloging at LC, for whom speed of “throughput time” was the consideration that trumped all others.)

Why is it so inappropriate for research libraries to regard speed in cataloging books as “the gold standard of quality service”? It is because scholars in research libraries do not need LC or any other library if their purpose is merely to find information quickly. Nor will they turn to books to begin with if speed is their dominant concern. They can already find “something” very speedily through Google, Yahoo, AltaVista, Amazon, or any of a multitude of other Internet alternatives—as well as through the hundreds of commercial databases to which their research libraries subscribe. Anyone who needs information immediately will generally be much better served, right from the start, by formats other than books—i.e., Web sites, newsletters, periodicals, encyclopedia articles, online abstracting services, newswires, or broadcast transcripts.

Here is the point utterly missed by Calhoun: Books do not yield their insights through fast keyword searching—they usually have to be read through for their content to be grasped. And that is not usually the kind of format being sought, to begin with, by people who want “something quickly.”

Books, by their very nature, are more important to scholarship than to quick information seeking. And scholarship is directly and severely undercut, rather than enhanced, if “getting the books on the shelf quickly” also entails the dismantling of the cataloging structures that provide the best access to them—structures that collocate relevant texts (no matter what their keywords or languages) so that the “right” words within them appear only within the desired, and limited, intellectual contexts; that systematically map out unanticipated aspects, connections, and linkages in ways that enable researchers to recognize relevant options that they cannot specify in advance; and that integrate all works, past and present—not just the most recent, or only those
currently in print—within the same conceptual categorizations. The Calhoun report shows no awareness at all of the requirements of scholarship—it dumbs down all research to the level of quick information seeking that is partial, haphazard, and “hit and miss.”

Structures that bring about systematic retrieval cannot be created quickly by technicians, by computer algorithms, by “under the hood” programming behind a blank keyword search box, or by “throwing more keywords into the hopper.” They can be created only by professional library catalogers—whose work AFSCME 2910 is proud to represent.

Skewed sample of interviewees

Yet another defect in the Calhoun report is the skewed sample of experts chosen to answer its questions. It is immediately obvious that the author sought out feedback from institutions heavily invested in the Google Book Search project (Harvard, Michigan, Columbia), from the California Digital Library, from OCLC, and, of course, from Sarah Thomas’s Cornell. While all of these worthies do indeed have a legitimate place at the discussion table, there is a rather serious problem with the omissions. It is noteworthy that neither OCLC’s nor Google’s software can display the left-anchored browse displays that are so important a feature of LCSH. Other stakeholders’ systems, however—RLG, EBSCO, H.W. Wilson Co.—have softwares that can and do display these browse displays, and they rely on the provision of linked strings supplied by LC. Why does this commissioned report lean so heavily towards the limitations of OCLC’s and Google’s capabilities to the detriment of their competitors in the market place, whose software makes much more efficient use of LCSH? What will happen to the “market position” of those competitors if LC suddenly undercuts them by failing to provide the linked strings of LCSH terms on which they depend? Should a federal agency such as LC be using its enormous influence to “tilt” the market toward OCLC and Google? This is the kind of thing that might well invite Congressional inquiry.

The linkage of Library of Congress Subject Headings and Library of Congress Classification

More important still: there are currently 379 paying subscribers to LCSH plus an additional 1,151 institutions that receive their copies through depository collections (total 1,530). Most of them, of course, also use the Library of Congress Classification system (LCC) for arranging their books on the shelves. LCC, however, is intimately intertwined with the LCSH subject terms that serve researchers as its index. (Look at any page of the LCSH volumes to see how many headings are directly tied to how many specific LCC classification areas.) If research libraries no longer have LCSH, then they will effectively have no way to point users to which areas of the bookstacks they need to browse for the clusters of books on their topics—and, of course, the catalogers’ ability to make systematic class designations in the first place will also be severely diminished. (Note from the Appendix on User Studies, below, how important browsing in the stacks continues to be even in this “digital age.”) Calhoun, of course, shows no awareness at all of the importance of this linkage of LCSH and LCC—or even of its existence. Her report
suggests that LCSH can be simply abandoned with no repercussions on how research libraries determine which classes their books should be shelved in. (The term “bull in a china shop” tends to suggest itself at this point.)

**Onsite book collections to be replaced by shared warehouses**

But then Calhoun also implies that onsite book collections shelved in LCC order are not necessary to begin with! As she says, “A lot of funds are currently locked up in building many parallel, redundant research library collections. Therefore, a first step in implementing the leadership strategy must be helping libraries pool their collection efforts, freeing up materials budget funds for reinvestment” (p. 15). On the same page she takes the Association of Research Libraries to task because “The structure of most ARL budgets privileges the purchase of printed books and serials for locally-housed, locally-circulated collections. Research libraries continue to be ranked primarily by how many things they have in locally-housed, locally-circulated collections.” The one footnote she supplies to justify this disdain for local collections is a paper that calls for shared warehouses in preference to local collections—and warehouse storage, of course, will done according to bar codes and books sizes, not according to the Library of Congress Classification scheme. The italicized words are all terms of opprobrium for Calhoun.

In short: If it’s not digital, if it’s not remotely accessible online [what happened to Copyright?], if it involves professional human judgement or expertise, if it entails conceptual categorization rather than relevance ranking of keywords, then it’s a library operation that just doesn’t fit the “leadership” model of the Calhoun report.

**The Bed of Procrustes**

There is a famous story in Greek mythology about the character Procrustes; *The Oxford Classical Dictionary* says of him, “Having overcome his victims he would force them to lie down on a bed, or on one of two beds; if they were too short, he would hammer them out or rack them with weights to fit the longer bed, if too tall he would cut them to fit the shorter.” The Calhoun report is very much a latter-day Procrustean bed for the library field. It forces academic institutions into an inappropriate business model, and lops off the goal of scholarship if it does not fit the criterion of increasing “market position.”. It distorts that business model itself. It pulls unjustified conclusions from the misuse of a biological metaphor. It inflates wishful thinking on the acceptability of e-books to the level of established fact. It lops off any consideration of the “niche” model that research libraries should focus on promoting scholarship rather than quick information seeking. It *reduces* the needs of scholars to the immediate wants of “information seekers.” It lops off any consideration of the need to systematically retrieve foreign language works in the same sets as English language works. It utterly disregards the scholarly advantages of LCSH. It grossly inflates the utility of relevance ranked keyword searching. It lops off any considerations of the system of elaborate linkages between LCSH and LCC. It minimizes the importance of research libraries having “redundant” onsite book collections for their local scholarly communities. It inflates the importance of “fast turnaround” while mistaking the
distinctive role of *books* in the scholarly research process. It represents the views a highly-
skewed selection of interviewees as representative of the larger library community. It involves a
federal agency in tilting the library market in favor of some stakeholders to the detriment of
others. And it simply ignores the views of hundreds of subscribers to LCSH. Its defective
scholarship and heavily “agenda’d” perspective do not form a sufficient basis for the Library of
Congress, or the larger research library community, to accept it major recommendations.

Finally, one more point must be raised; it is a question that may legitimately be asked of a
public, taxpayer-supported institution: How much money did LC pay for commissioning this
study?

One wonders how committed Calhoun herself is to a business model, for in most such
arrangements a “product” with as many defects as the Calhoun report can be returned for a full
refund.

Appendix

Survey of Library User Studies

1) Council on Library and Information Resources (CLIR) and the Digital Library Federation,
*Dimensions and Use of the Scholarly Information Environment*. Online at <
http://www.clir.org/pubs/reports/pub110/contents.html> (See Table 650.)
This study found, based on over 3,200 interviews, that 55.4% of all respondents (and 59.7% of
undergraduates) still regard browsing library bookstacks as “an important way” to get
information. **Two thirds of the faculty and grad students use print resources for research
all or most of the time (73% for teaching). 52% of undergrads use print resources for
coursework all or most of the time.** (The figure was 72% for grad students.) More than
90% of them agreed that print books and journals “will continue to be important sources for me
in the next five years.” . . . **59% of undergrads still use print abstracts and indexes, 93% use
printed books, and 81% use print journals (97% of grad students use print journals). Only
28% said they “find reading information on a screen satisfactory.”** 86% of students feel that
“my campus library meets most of my information needs.” **55% still regard browsing the
stacks and journal shelves as an important way to get information—and only 35% use the
library significantly less than they did two years ago.** 14% want more print journals as
compared to 11% who want more e-journals—and 89% want more books.

2) College Student Experiences Questionnaire, with data representing “responses from more than
This survey found that 65.5% of male students, and 63.2% of female students, reported that they “found something interesting [through] browsing” either “occasionally,” “often,” or “very often.” For students in “Doctoral Intensive” programs the overall percent is 67.7%, with 24.9% reporting that browsing [in library bookstacks] was useful either “often” or “very often.”

(3) **OCLC White Paper on the Information Habits of College Students (June 2002)** with sample of 1,050 qualified respondents. (Available at <www5.oclc.org/downloads/community/informationhabits.pdf>)

Among its findings are these:

- More than 31% of all respondents use Internet search engines to find answers to their questions. However, **people who use Internet search engines express frustration because they estimate half of their searches are unsuccessful.**

- Americans have not yet found an ideal information resources. Not one participant said they would use the same resource time and again when seeking answers.

- Nearly 9 or of 10 students (89%) also use the campus library’s print resources, including books, journals, articles, and encyclopedias.


This inquiry, which interviewed faculty at Oklahoma and also surveyed other ARL libraries, notes the following:

“Part of the faculty space’s value to researchers lies in its proximity to the collections of monographs and print journals. **The importance of serendipitous browsing in library collections cannot be overemphasized by the majority of faculty space holders.** ‘It’s that one minute out of fifty-nine [minutes], when you find that one gem on the shelf’ that makes the act of browsing not only effective but absolutely vital to many researchers. As one participant emphatically noted, ‘There is no substitute for walking the stacks. It’s not “browsing”—that sounds too aimless. It’s more directed—‘surveillance,’ really.’ . . .

During the interviews, **the theme of serendipitous browsing emerged repeatedly with regard to research.**”

(5) The Electronic Publishing Initiative at Columbia (EPIC) survey of responses from 1,233 students and
Direct quotations from various passages of the “Survey of College Students” section:

• “Students are almost as dependent on the physical library (75.8%) as they are on the library’s website to retrieve books and articles (81.5%).”

• “Books and journals are still cited by most students when writing a term paper, however the number of students citing websites does not lag far behind. Books and journals were each cited by more than three-quarters of respondents as types of resources cited in the bibliography of their last research/term paper (84.8% and 77.8% respectively), while websites were cited by 68.8% of the students.”

• “[F]or general assignments, students are more likely to turn to the Internet, but for in-depth research assignments, they are somewhat more likely to turn to library sponsored [i.e., restricted access] electronic resources.”

• “Undergraduate students are more likely than graduate students to use non-library sponsored electronic resources, while graduate students are more likely to use library sponsored electronic and library sponsored print resources.”

• “The majority of respondents use the physical library more than once a month (67.7%).”

• “Print is preferred for situations where the material is long or dense, and the reader has to fully comprehend the material. Electronic resources are preferred for situations where the reader is obtaining supplementary or background materials, for current events materials, or for looking up information for short papers/homework assignments” [emphasis added].

• “The physical library is still an important destination for students.”

• “Search engines appeared to be a major difficulty in the use of electronic resources. 80.2% of respondents agreed that search engines were not as precise as they would like them to be and 67.3% agreed that search engines are not as thorough as they would like them to be” [emphasis added]. [Note that these statistics on the extent of the problem are conspicuously absent from the report’s “Executive Summary” of its Survey of College Students.]

• “Further, 71% agree that electronic resources increase their need to separate out the reliable from unreliable information, while half of the respondents report difficulty making these judgments.”

• “More than half of the respondents somewhat strongly agree that electronic resources can result in an overload of information . . . and almost half of the respondents agree that this overload can be overwhelming for them.”

- “Internet use does not reduce library use.” [emphasis in original]

- “In [this] study, we found that while they tended to use the Internet to get news, health information, recipes and other short-term’ material in a brief format, they used the library for in-depth research and extensive reading.”

- “They reported that 75.2 percent of Internet users also used the library, 60.3 percent of library users also used the Internet.”

- “They found no evidence among respondents who used both the library and the Internet that Internet use was changing the reasons why people used the library or the frequency of their library use, but that respondents used each information source for different reasons.”

7) “Historians and Their Information Sources,” a 2003 survey of 278 university-situated historians. (Margaret Steig Dalton and Laurie Charnigo. “Historians and Their Information Sources.” College & Research Libraries 65 (September, 2004), 400-425.)

- Table 1: Materials Considered Important for Research

<table>
<thead>
<tr>
<th>Materials</th>
<th>Total [responses]</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Books</td>
<td>274</td>
<td>99%</td>
</tr>
<tr>
<td>Journal articles</td>
<td>273</td>
<td>98%</td>
</tr>
<tr>
<td>Web sites</td>
<td>162</td>
<td>58%</td>
</tr>
<tr>
<td>* * *</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Table 4: Most Frequent Ways of Discovering Primary Information

<table>
<thead>
<tr>
<th>Method of Discovery</th>
<th>Total</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finding aids</td>
<td>64</td>
<td>26%</td>
</tr>
<tr>
<td>Footnotes</td>
<td>50</td>
<td>20%</td>
</tr>
<tr>
<td>Archival/library catalogs</td>
<td>49</td>
<td>20%</td>
</tr>
<tr>
<td>Web sites</td>
<td>7</td>
<td>3%</td>
</tr>
<tr>
<td>* * *</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- “Systematic bibliographic searches in databases, use of other sources such as aids designed to disseminate information about resources, and careful, patient, wide reading followed by pursuit of leads resulting from the reading are clearly the foundation of historical research; the package was summarized by one historian as ‘diligence.’” [p. 409]

- “One follow-up question asked the historians whether quality, availability, or ease of use most affected their choice of an information source. Most reported that quality was the most important
consideration . . . The historians, however did seem willing to exert themselves to obtain what they regarded as necessary.” [p. 410]

- “Serendipity . . . plays a significant role in historical research. Browsing is an invitation to opportunity and was a frequent method of choice. Browsing usually meant scanning library shelves in an area that was a logical, but not guaranteed location of information . . . many testified to its value.” [p. 410]

- “Comprehensiveness is clearly the highest priority in searching a database.” [p. 411]

- “Almost two-thirds (64%) used the Internet from home, and responses to the question about which Web sites were most often visited showed considerable familiarity with the range of possibilities. . . . The single most frequently visited Web site listed was that of the Library of Congress” [pp. 412-13]

- Conclusions
  “Comparing the results of this study with those of the 1981 Stieg study demonstrates that although much related to historians’ information-seeking habits has changed, many things have remained the same. . . . Browsing is still important. Print remains the principle format of the information used, although electronic databases are used extensively in the discovery of information, and books still dominate the discipline. This domination of print is even clearer in the citation analysis than in the survey results. . . One important change that has taken place is that catalog and index use has grown and continues to grow.” [p. 417]

8) “The Internet and Education: Findings of the Pew Internet & American Life Project,” a “survey of 754 children, ages 12-17, who use the Internet and one of their parents or guardians.”

According to this study, “When asked to think about the last big report they wrote for school, 71% of online teens reported relying mostly on Internet sources for their research. Another quarter (24%) reported using mostly library sources, and 4% said they used both equally.”

Some reasonable qualifications must be noted:

1) This study is a survey of children below college age; the survey itself uses the word “children” (or “teens”) to describe them. The numerous surveys of college undergraduates, grad students, and senior scholars (above) report much higher percentages of library use. It is not at all surprising that children’s grade school and high school homework assignments do not require in-depth or substantive scholarly research.
2) The wording of the findings, even in this study itself, is hedged: while it says that 71% of these children relied “mostly” on Internet sources, it does not assert that even this group used *only* Net sites.

3) The study does assert that “Students cite the ease and speed of online research as their main reasons for relying on the Web instead of the library,” and, indeed, the phrase “instead of” (not “in addition to”) may well be accurate regarding the majority of these underage respondents. The limited applicability of this finding, however, based as it is on a survey of an immature population that is required to do only superficial research, must be viewed in the context of the many other surveys that consistently report much higher percentages of library use by older researchers who are required to do research that is both more substantive and more comprehensive.

4) Although the study also interviewed the parents or guardians of these 754 children, and while it did ask these adults about their own use of the Internet, it did not ask the adults any questions about their own use of *libraries*. Nor does the study suggest that any of these adults have particularly academic interests or job requirements.

In sum, it is a distortion of the evidence to assert that the Pew study “proves” (or even “shows”) that “most students do not use libraries.” One would have to equate college students, graduate students, and senior scholars (as well as the general adult population) with underage children in order to hold such a belief.

9) “Counting on the Internet: Most Expect to Find Key Information Online, Most Find the Information They Seek, Many Now Turn to the Internet First,” from the Pew Internet & American Life Project (2002). (Available at: <http://www.pewinternet.org/report_display.asp?r=80> (or combine “counting on the Internet” and “Pew Internet & American Life” in a Net search engine)

This study confirms that many Americans now use the Internet as their first source in seeking information; but it says nothing at all about library use. It does not even mention the words “library” or “libraries.” From several of the other studies, above, that did ask about library use, we know that most researchers do not confine their efforts solely to the Internet. It is a misrepresentation of this Pew study to claim that it indicates people do not use libraries. It is silent on this question—it did not ask respondents about their library use.