

Purdue University
Purdue e-Pubs

Charleston Library Conference

“Eat Yourself Full, Leave Your Plate Empty”: Or Why Student and Faculty Appetite for Data Is Like an Offensive Lineman at a Buffet

Angela K. Horne
University of California, Los Angeles, ahorne@library.ucla.edu

Corey Seeman
University of Michigan, cseeman@umich.edu

Rebecca A. Smith
University of Illinois at Urbana-Champaign, becky@illinois.edu

Follow this and additional works at: <https://docs.lib.purdue.edu/charleston>



Part of the [Library and Information Science Commons](#)

An indexed, print copy of the Proceedings is also available for purchase at:

<http://www.thepress.purdue.edu/series/charleston>.

You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: <http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences>.

Angela K. Horne, Corey Seeman, and Rebecca A. Smith, “Eat Yourself Full, Leave Your Plate Empty”: Or Why Student and Faculty Appetite for Data Is Like an Offensive Lineman at a Buffet” (2013). *Proceedings of the Charleston Library Conference*.
<http://dx.doi.org/10.5703/1288284315269>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

“Eat Yourself Full, Leave Your Plate Empty”: Or Why Student and Faculty Appetite for Data Is Like an Offensive Lineman at a Buffet

Angela K. Horne, Head, Rosenfeld Management Library, Anderson School of Management, University of California, Los Angeles

Corey Seeman, Director, Kresge Business Administration Library, Stephen M. Ross School of Business, University of Michigan

Rebecca A. Smith, Business Librarian, College of Business Digital Research Library, University of Illinois at Urbana-Champaign

Abstract

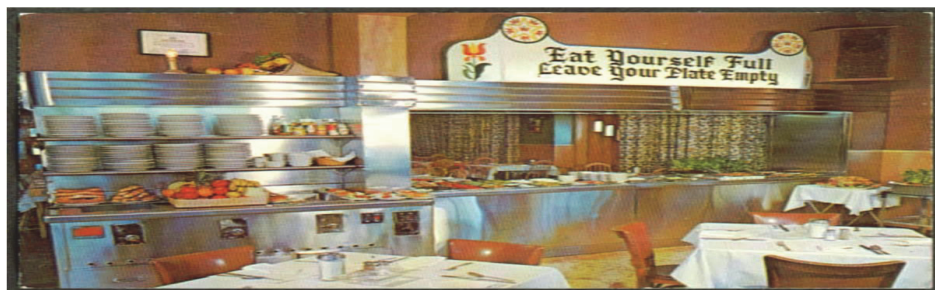
From an old postcard of Miller's Smorgasbord (on the Lincoln Highway east of Lancaster, Pennsylvania) we see the sign that stands over the scrumptious buffet welcoming all patrons at the restaurant: “Eat Yourself Full, Leave Your Plate Empty.” The notion is simple, take what you can eat but do not waste food. But in many ways, the whole premise of a buffet is the ability to try, sample, nibble, and experiment with foods that you might not order otherwise order. And we all pay the same, even if we are a college football offensive lineman with a legendary appetite.

It is this conundrum that leads us into problems between libraries and the database vendors. Librarians scramble to keep the balance between the students' needs of getting and analyzing data with the vendors' needs to keep the systems working and not to be overburdened with royalty payments to the publishers. Increasingly, we see faculty and students wanting to download more data, but vendors too quick to install restrictions out of intellectual property concerns. In this “Eat Yourself Full” database environment, our students and faculty want to download more and more to analyze and interpret on their own.

The presenters will explore the issues of downloading caps and other obstacles at the business libraries of three large U.S. public universities (Illinois, Michigan, and UCLA). Among the topics discussed will be the topics that drive this research, the types of resources they wish to use, the impact of the limits imposed on the students and faculty, and the workarounds that connected the user to the data that they needed.

Much has been written about buffet pricing and consumption. Brian Wansink (Cornell University) has conducted many studies of consumers at all-you-can-eat Chinese Buffets. In these observations of consumers, the studies looked at the estimated Body Mass Index (BMI) along with observed behaviors at the restaurants. These include: size of the plate used (large vs. small), facing the buffet or

facing away, sitting at a table or at a booth, using chopsticks or a fork, and browsing the buffet first or serving immediately. None of Wansink's studies dwell on the quality or taste of the food, but they did record the leftovers on the plate (ranging from 6% to 10% of food content [Wansink & Payne, 2008; Wansink & Shimizu, 2013; Wansink, 2010]).



How might these studies apply to library usage and downloading behavior? Students and faculty will download far more than they will use—and potentially read. Additionally, usage counts are driven every time we download a document and many are not fully “digested.” Like with groceries, we do not simply buy what we need, but buy in the quantities that are available. So while business librarians are being inundated with usage counts that seem too good to be true and restrictions or limited on usage and downloads, we thought it would be good to explore how business research is being done at three large public universities with top business programs. We looked at business research at UCLA, Michigan, and Illinois; all three libraries have members in the Academic Business Library Directors (ABLD). The ABLD represents 50 of the top academic business libraries in North America.

The key elements for each of these three schools:

- University of California, Los Angeles
 - Multiple funding streams (Anderson School of Management, UCLA Library)
 - Ceased active monograph purchasing
 - Many nonbusiness students and faculty on campus want our resources
 - State funding decrease: 70% (1970s) to <10% (2013). Move to self-sustaining.
 - Population Served:
 - Full-time MBA: 750
 - FEMBA, including Flex: 850
 - EMBA: 500
 - GEMBA: two programs
 - MFE: 100
 - PhD: 80
 - Undergraduate Accounting Minors: 200

- Faculty, including lecturers and adjuncts: 140
- Sample UCLA Research Topics
 - Rental rates by zip code
 - Regional sports data (ad spend, stadium concession sales)
 - Stock data for all companies across multiple exchanges
- University of Michigan (Ross School of Business)
 - Funding of library purchases (Ross School of Business)
 - Essentially flat funding last 5 years, but decreases in materials.
 - This year, large cuts in periodicals and standing orders in preparation of significant space constraints.
 - Kresge supports business research for departments outside Ross
 - Population served (Fall 2013)
 - Full-Time MBA: 941
 - Part-Time MBA: 315
 - Weekend MBA: 164
 - Global MBA: 38
 - Exec MBA: 196
 - Other Masters: 124
 - BBA: 1422
 - Undergraduate Minor: 100
 - PhD: 95
 - Faculty: 210
 - Sample University of Michigan research topics:

- Assessment of all companies in a field (financial performance)
 - Analysis of corporate press releases and news articles on a particular topic or trend
 - Determination of Market Sizes, especially in growing and developing industries (historical)
 - When the job market improves, reference demands decrease (and vice versa)
- University of Illinois
 - Funds for database are collaborative; some are interdisciplinary and paid by the ER fund; some for business library budget (University Library subject funds); most financial data sets are paid by business school but business librarians provide research support in their use
 - Population Served:
 - Undergraduates: 2,974
 - Master's Programs: 712
 - MBA: 296
 - Exe MBA: 56
 - PhD: 100
 - Faculty (FTE Tenure-Track): 109
 - Faculty, other (visiting, post doc, adjuncts, lecturers):90
 - Sample University of Illinois research topics and trends
 - In academic business research, faculty and PhDs want full access and more than what the buffet allows for each meal now that so many external drives can hold more memory
- They will sort out what they do not want later and/or add to what they have already downloaded for perceived time savings
- Some will distill using tools such as STATA, SPSS, SAS, or BibTex, Invivo
- Undergraduates, Walk-ins will turn to Google first, hoping the information needed is freely available, spend a lot of time that, in sum, is more than they anticipated
- Locate data sets or market research reports which are not free, then ask the library (via a chat service or e-mail sometimes in person) if the library owns the material
- The other scenario is that the student may be aware of market research reports and through topics or keyword, they find reports, download more reports than necessary, then sift through the reports for the two pages of data/text to support their research

In many regards, the issues that are ever present at business libraries is balancing between the needs of the vendors and publishers with the desires for information from the students and faculty. The research trends at these three schools

demonstrate a tremendous appetite for data and information that some vendors are not comfortable with. We have talked with vendors who have told us that students can download as much as they want. But in pressing them, we discover that there are limits that most transactions (over 99%) will not run into. That being said, we find many instances where business research pushes past these limits as students attempt to study trends and make value determinations for assignments. This is especially true of faculty and PhD students who often require more data than can easily be extracted from an online resource.

Issues with Vendors and Libraries

The speakers explored some of the issues with vendors recently that were of particular interest for business libraries. These include: Harvard Business Review and EBSCO access (Business Source Complete), Financial Times and 30-Day embargo, data mining and Factiva, PrivCo access and usage problems, problematic audiences, and perceptions of misuse by academic community members.

Harvard Business Review (HBR) is one of the leading publications in the business space and a true standard-bearer in EBSCO's Business Source Complete (BSC). EBSCO is the only online version for Harvard Business Review, and it includes the entire run from 1922 to the present in full-text format. While users had full-text access to these articles, use in the classroom was never implied in this access as it was only for personal research use. Harvard Business Publishing viewed HBR articles similarly to cases, which do not have fair-use rights in the classroom. Any classroom use required payment of rights through Harvard Business Publishing.

During summer 2013, a significant change took place in that Harvard selected their top 500 articles (based on usage) and made them view-only via EBSCO. If a library wanted to lift that restriction, their only solution for Harvard access to get full download access was to acquire their Enterprise License. At the University of Michigan, where the Kresge Library manages a course pack operation, they have had the enterprise license in

place for a few years, so their access was not affected. The cost for the enterprise license for schools ranged a great deal and seemed to be driven by the number of expected uses in course materials during the year.

While there has been a large backlash against Harvard in this regard, we do not see any change coming down the pike. In many ways, it seems that HBR might not be long for an aggregator database given the usage restrictions. But time will tell there.

Also during summer 2013, the Financial Times, a leading business newspaper published globally from London, changed the way that universities had access to their content. The content, which is available via numerous aggregators including ProQuest, LexisNexis Academic, Factiva, and others, was changed to incorporate a 30-day embargo for new content. Nearly simultaneously, they introduced a new campus access program which would provide access to the most recent 30 days, plus some additional content. The significant cost for the Financial Times was not possible for the three presenting libraries to absorb. We are watching for possible others following this path, such as the New York Times as they are launching a similar service for libraries.

With Factiva, the University of Michigan was caught in a problem involving excessive use from a student. The student was conducting research, but the quantity was considered to be a data-mining exercise which is against the standard contract. The standard contract includes this prohibited activity: "Use of Information or the attached codes to feed any data-mining software or other automated trend analysis application." The librarians at Michigan were unaware of this research, and the IP that was showing on their servers was the campus IP number. Factiva slowed down access and downloading until the usage spike stopped. We were able to find out who was doing the research because of questions asked of another department. During our "breach," Factiva set a limit of one download of an article at a time (as opposed to 100). In order to do the work that the student desired, we would need to purchase a "text mining license." We were able to avoid this repeating itself because a subsequent researcher just focused on using Boolean searches

and counting results that resulted in no spike because there were no downloads.

Another database where problems arose was PrivCo, a new database that provides information on privately held companies. Information on privately held companies is hard to come by. It should come as a little shock to anyone that these are private for a reason. Whereas public companies need to share a great deal of information (and, therefore, the data are abundant), private companies remain one of the most requested topic by researchers at business libraries. PrivCo has quotas on downloading, but they did not explain that in the license agreement that they signed with at least one ABLD Library (Illinois). They have not made it clear when the quota limit hits, but it is known that the download limit is about 900 lines at a time in Excel format. In mid-April 2013, the business librarian at Illinois received an e-mail from a sales representative that they were concerned about too much downloading and they were cutting off access for 2 weeks until they could supply a more workable situation (this also happened at another ABLD library). Illinois pushed back because of timing in the semester (3 weeks until finals) and that students should not be penalized for accessing the database when papers/projects were due. In a compromise to get through the remainder of the semester, PrivCo would mediate search of companies and their profiles and send them to the students (this had been done previously for PhDs/faculty) who requested them. While the quota issue remained unresolved, access was restored within 2 days. This is one model that is okay for PhDs and faculty who are comfortable with the arrangement as long as the data are delivered within a few days, but not sustainable for undergraduate students or MBA students who want a few companies' data tomorrow for project analysis. This remains a current issue with the library, and we are continuing to explore options with the company.

Another problematic area for business libraries is with the audiences that we serve. Academic customers pay "pennies on the dollar" for most business content versus what commercial customers pay, and some vendors are very

concerned about how this might be abused. The biggest issue is with part-time MBA students, be they in Executive MBA programs or part-time MBA programs. The notion is that while they are a student, they have access to a large number of resources that they might use for their work. While many libraries will tell students that that is not permitted via most contracts, this is nearly impossible to police for the library. One vendor (Capital IQ) has strong restrictions about how the products may be used, and they make great strides to limit the perception or potential of abuse, including turning off access during the summer. Additionally, there are many products that specifically prohibit walk-in usage or alumni use. These are more easy to control, but may fly in the face of many desired policies at the library. And to this end, many libraries have been confronted by vendors because of a perceived misuse of the data. Some vendors have asked that we put language to the effect that these are for academic use only on some of the databases. However, there is really no way to police that. This is where we often find that we run into problems balancing between the needs of our community and the vendors.

Solution, Ideas, and Recommendations

Much of the problems that we face in this space are what we refer to as the problems with life at "the knot." In many regards, librarians and vendors work out arrangements, but neither are the users or content creators. For many vendors, such as ProQuest, EBSCO, or Gale, they are licensing content from publishers. And for many publishers, they are setting rules requested by the organizations and learned societies who are publishing through them. In the library space, it is not librarians who are using the content, but the students and faculty. So this is a situation where the needs of the content producer and the end users are difficult to manage when the license is between two intermediaries. The ABLD libraries are particularly aware of the need to balance between the needs of the users and the needs of the vendor. This stems from the fairly unique issues associated with business research, especially when there are distinct pricing models for the data and resources in the academic and

commercial marketplaces. Academic customers routinely pay pennies on the dollar, in large part, to get in front of their future customers.

Another potential solution that has been implemented at a number of schools is to work with vendors that do not want unmediated access to their resources. There is a role for mediated searching and/or downloading by librarians on behalf of students. Many vendors are nervous about offering access to the end user, but are okay with a mediated approach. One example at Michigan is with Frost and Sullivan, an excellent market research firm that produces reports that are highly sought after by Ross students. With Frost.com, the generic log in (used by students) does not allow downloads of reports. However, they are told to contact a librarian with the report information, and the librarians can download it for the students. This provides the students with the report and the company with the security they desire. This also provides publishers with the “breaks” on runaway usage (a real threat for them). Contrary to what one might think, there is not a ton of work here for the librarians. This is a nicer balance to vendors who do not allow for downloading of reports for academic customers.

Downloading large data sets is a very common need that students have. This might be related to large batches of company records or news articles for analysis. Many resources have smaller caps (25 or 50 up to 5,000) that do not work for all

projects. These caps are put in place to both protect the content and to ensure that no request bogs down the entire system. While we can understand the needs for caps, our students have research projects that simply cannot be done within the confines that are needed by the publisher. One vendor (Bureau van Dijk) has been working with libraries to sell them a snapshot of their Orbis data to enable this type of downloading. Orbis is the largest company database available and includes over 100 million business records worldwide. The records are delivered to us on a Blu-ray Disc, and we are building a small library. This allows the students and researchers to download whatever data they need without taxing their web system.

Other issues that are faced by business librarians include:

- Using “list price” of market research reports to imply great value when we pay only a small portion of that price in the academic space. This resembles a sale at Kohl’s.
- Too often, the data requests stem from individual research requests, not longstanding research needs. We are being asked to buy when we really only need the information for a short time. It would be akin to buying a car when you travel, not rent one.

References

- Flores, D. (2009). All you can drink: should we worry about quality? *Journal of Regulatory Economics*, 35(1), 1–18.
- Just, D. R., & Wansink, B. (2011). The flat-rate pricing paradox: conflicting effects of “all-you-can-eat” buffet pricing. *The Review of Economics and Statistics*, 93(1), 193–200.
- Nahata, B., Ostaszewski, K., & Sahoo, P. (1999). Buffet pricing. *The Journal of Business*, 72(2), 215–228.
- Wansink, B. (2010). From mindless eating to mindlessly eating better. *Physiology & Behavior*, 100(5), 454–463.
- Wansink, B., & Payne, C. R. (2008). Eating behavior and obesity at Chinese buffets. *Obesity*, 16(8), 1957–1960.
- Wansink, B., & Shimizu, M. (2013). Eating behaviors and the number of buffet trips: An observational study at all-you-can-eat Chinese restaurants. *American Journal of Preventive Medicine*, 44(4), e49–50.
- Wu, D., & Liu, N.-P. (2007). Why do amusement parks only charge a fixed admission fee? *Economics Letters*, 95(2), 180–185.