UC Berkeley Library Licensed Electronic Resources: Description and Analysis

David Moles

School of Communication and Information, Rutgers University

17:610:532:90: Collection Development and Management

Dr. Stephanie Mikitish

April 17, 2023

UC Berkeley Library Licensed Electronic Resources: Description and Analysis

Introduction

The University of California, Berkeley (hereafter, UC Berkeley) is a public, coeducational research university, offering a variety of programs from bachelors' degrees through research and professional doctorates. The university was founded in 1868, as a merger between the private College of California and the state land-grant Agricultural, Mining, and Mechanical Arts College; it moved from Oakland to Berkeley in 1873 (UCB, n.d.). It is the oldest and, by number of students, the second largest (after UCLA) of the ten campuses of the University of California system (UC 2022), serving more than 45,000 students, roughly 30% graduate and 70% undergraduate (UD&AR, 2022, p. 5). UC Berkeley prides itself on its research, advertising the facts that 93 Berkeley students received NSF Graduate Research Fellowships in 2022 (p. 9) and that, among other honors, the faculty has included 26 Nobel Prize winners, ten currently teaching (p. 5); and counting Harvard, MIT, Stanford, Oxford, Yale, and Princeton as peer institutions (pp. 5, 9, 11).

The UC Berkeley Library is the university's main library organization, comprising 23 campus libraries holding more than 13.5 million volumes (UCB Library, 2022), including more than 2 million e-books (UCB Library, 2021). There are also eight affiliated libraries, including those associated with the law school, with various academic disciplines, and with the Berkeley Art Museum and Pacific Film Archive (UCB Library, 2022). The UC Berkeley Library operates as part of a consortium with the other nine UC campuses as well as the California Digital Library (CDL), attached to the UC Office of the President; in 2021, after four years of work, the ten campus library systems were consolidated in a single systemwide ILS (SILS) based on Ex Libris' Alma ILS and Primo VE discovery system (UC Libraries, n.d.; CDL, n.d.). Full library privileges, including access to electronic resources, are available to students, staff, and faculty, as well as to visiting scholars, postdoctoral scholars, and retired faculty and staff, as well as to some staff from affiliated agencies such as the UC Office of the President whose primary job includes

working on behalf of UC Berkeley (UC Berkeley Library, n.d.). Most external users do not have access to licensed electronic resources, but access to most electronic resources is available to the public on a walk-in basis, via public computers in library reading rooms.

Licensed Electronic Resources at the UC Berkeley Library

For this project, I will be examining the UC Berkeley Library's collection of licensed electronic resources, including databases, journal packages, and individual journal subscriptions. This excludes those consortial resources licensed separately by CDL, and those resources managed by CDL though licensed by an individual campus (including Berkeley) for multi-campus use; but while the quantity of content licensed consortially is rather larger than that licensed by the Library exclusively for use by UC Berkeley patrons,[1] the latter is still considerable. According to data provided to me by Systems and Discovery Services Librarian Jackie Gosselar (personal communication, March 27, 2023), UC Berkeley-licensed resources comprise some 1400 packages and databases and more than 81,000 journal titles, some licensed individually and others via packages, which in March 2023 saw roughly 1.2 million downloads.[2] Assuming an annual rate of perhaps 12 million downloads (allowing for reduced summer usage), this compares favorably with the 6.8 million downloads by UC Berkeley patrons of CDL-managed consortial resources (CDL, 2019), suggesting that on the whole the collection is meeting the needs of its users. The most popular 0.1% of titles accounted for more than half of all downloads;[3] more than 90% of titles had no downloads whatsoever during this period. The

---

[1] I was unable to find any official CDL figures, but according to the LinkedIn page of CDL Assistant Director of Systemwide Licensing and Collection Services Lisa Mackinder, CDL licensed content is "worth in excess of $45 million" (Mackinder, n.d.), as compared to UC Berkeley's annual budget for licenses and subscriptions (including print subscriptions) of $13.8 million (USDOE, n.d.).

[2] Estimate based on total of 984,205 downloads from March 1 to March 26; but see note to Table 1.

[3] In the period covered by the data, 50.1% of downloads (492,809 of 984,205) were of the 76 most popular titles, representing 0.094% of 81,070 titles in total.

most popular title in March 2023 (see Table 1) was *Vogue*, with nearly 10% of total downloads; the *Journal of the American Chemical Society*, with 5%, was in second place.

Selection and Deselection Criteria

In general, the UC Berkeley Library prefers electronic access for serials when possible, due both to its relative ease of access and discovery, and to the increasing costs of print retention (Newyear Ramirez & Silva, n.d.). Specific titles should be relevant to current teaching, curriculum, and research, or otherwise meet the unique needs of a department, such as accreditation or degree certification. Resources should be discoverable within the SILS Primo discovery system, and available to patrons based on access from UC Berkeley IP addresses, without passwords or individual patron accounts.[4] Deselection criteria include declining or low usage, unusually high cost per use, unusually high price increases, and duplication—whether across packages, or of content available through systemwide licensing and other consortial agreements, or of content available via open access. Where possible, the policy of the UC Libraries and the systemwide Faculty Senate is to prioritize investing in open access over paying for gated content, in the interests of more equitable access to knowledge, and of expanding the impact of UC generated research.

Proposed Analysis

Historically, in analyzing e-resource and other serials collections, the UC Berkeley Library has largely relied on what Johnson would describe as use studies (2018, pp. 302-302), cost-benefit analysis (p. 292), and/or collection profiling (p. 296), primarily with the goals of identifying low-usage titles, titles and packages with high cost per use, and duplication of titles across packages or between titles/packages licensed by the Library and those available

---

[4] Though the rationale for this policy is not given explicitly in the Library's internal documentation, I would assume it is not to only protect patron privacy, but also to allow walk-in access from Library reading room computers by members of the public.

elsewhere. Such assessments have been project-based and ad-hoc rather than continuous and formalized, driven by successive waves of budget reductions: student enrollment at UC Berkeley is 40% higher than in 2003, while Library funding per student, adjusted for inflation, has fallen by 47% (UCB Library Communications, 2022); as a result, since 2017, the Library has implemented three rounds of cuts to collection budgets (UCB Scholarly Resources, 2018; 2020; 2022), with the budget falling from $22.8 million in FY2017 to $19.5 million in FY2020 before rising to $21.7 million in FY2021: $13.8 million for serials and other subscriptions and $7.9 million for monographs and one-time purchases (USDOE, n.d.). The Library is currently in the second year of a two-year reduction process with the goal of cutting the collections budget by a total of $1.7 million from FY2021 levels (UCB Scholarly Resources, 2022), and has recently released a list of roughly 1200 titles proposed for cancellation in 2023, with an anticipated savings of roughly $1.2 million (internal communication, April 4, 2023).

I believe that the Library would benefit from a formalized continuous assessment process, involving not only the under-resourced Library Assessment Program (consisting of just one full-time librarian, plus occasional student workers on a project basis) but staff across the library as an ordinary part of their duties, along the lines of the collection management cycle at California State University San Marcos outlined in Lantzy et al. (2020). While usage statistics such as those summarized in Table 1 are relatively easy to gather, they capture only part of the collection's value, and can provide no guidance with regard to materials *not* held by the library that might be useful to collect; they also require a great deal of interpretation. It seems unlikely, for instance, that *Vogue* is in fact the publication most in demand for research and instruction; and that a publication in a particular narrow field of study has low or even zero usage in one time period—as do some 90% of UCB journals, in the period covered by this data—does not necessarily mean it will not be essential in the next. I would therefore propose two forms of assessment to be implemented on an ongoing basis, one quantitative, one qualitative: citation studies (Johnson, 2018, pp. 300-302) and focus groups (p. 305).

Citation Studies

As a research university, UC Berkeley has somewhat different library needs than CSU San Marcos, which is primarily a teaching institution. Citation studies, particularly the tracking of citations in research produced by the library's user community (Johnson, 2018, p. 301), can potentially assess the usage of publications in research directly, and with statistics better correlated to their value for that use than raw downloads; while including undergraduate theses and capstone projects, as in the projects by Kohn and Gordon and by Datig cited by Johnson, might provide a reasonable proxy for usage in teaching, especially when combined with syllabus analysis as reported by Lantzy et al. (p. 164). As Johnson notes, this type of citation study cannot identify gaps in the collection, resources that "the student would have used if he or she had access to more resources" (p. 301), but this could be somewhat mitigated by working with CDL to implement such citation tracking across the UC system. CDL's eScholarship system, UC's systemwide open access repository and scholarly publishing platform, offers an ideal corpus of publications for citation analysis and is the obvious point at which to address problems of citation extraction and standardization, while partnering with CDL and with other UC campuses would spread the (considerable) labor and cost of constructing and maintaining the necessary citation data system more widely. CDL also has its own journal impact research, as well as partnerships with other universities and consortia and with organizations such as COUNTER, which could inform the effort and might eventually allow extending the data collection and analysis beyond the UC system. Making citation analysis continuous and ongoing would not only amortize the cost of setting up systems to support that data collection and analysis, but also help mitigate the problem of time lag identified by Johnson (p. 302) and better help the Library identify the changes in the scholarly publication landscape easily missed by one-time studies.

Focus Groups

Citation studies nonetheless, as Johnson notes (2018, p. 301), cannot capture the value of materials not cited directly. And as with usage statistics, citation studies work best in those fields

with a large number of researchers and/or students. (Were the Library to reduce its holdings only to those required by the University's most popular programs, it seems unlikely that the University would long retain its status and reputation, and would quickly find difficulty attracting students and faculty even to those programs.) As a complement to citation studies, therefore, I would propose an ongoing series of focus groups (p. 305), leveraging the existing subject-matter expertise and faculty relationships of the Library's subject liaison librarians. Such focus groups would ideally include both research and teaching faculty, postdoctoral scholars and research staff, and students at all levels, in groups both mixed (to expose the participants to multiple perspectives) and demographically separated (to reduce the effect of power differentials on participants' willingness to express themselves).

Johnson (2018, p. 305) identifies several issues with focus groups; most are procedural—determining the number and constitution of groups, selecting and recruiting participants, ensuring good facilitation and note-keeping, and avoiding interviewer bias—and can, I submit, be overcome by a collaborative effort among subject liaison librarians to identify and implement best practices including, likely, training in facilitation, and sharing of labor across subject areas. As a conceptual issue, Johnson argues that it is difficult to measure focus group results objectively; but without getting into the question of whether objectivity is *possible,* I would nonetheless question whether, in this case, it is *desirable*. The value of the focus group seems to me to be precisely in capturing the messy, subjective input of both focus group members and the subject liaison librarians themselves.

References

California Digital Library. (n.d.). *Systemwide ILS (SILS)*. https://cdlib.org/services/d2d/sils/

California Digital Library. (2019). *E-Journal article downloads* [Excel spreadsheet; UC access]. https://sp.ucop.edu/sites/cdl/cpa/_layouts/15/DocIdRedir.aspx?ID=SPCDL-695059133-416

California Digital Library. (2023). *CDL licensed resources with tier info* [Excel spreadsheet; UC access]. Retrieved April 12, 2023, from https://sp.ucop.edu/sites/cdl/cpa/licensed-resources/CDL-licensed-resources-with-tier-info.xlsx

Johnson, P. (2018). *Fundamentals of collection development and management* (Fourth edition). ALA Editions.

Lantzy, T., Matlin, T., & Opdahl, J. (2020). Creating a library-wide collection management cycle: One academic library's approach to continuous collection assessment. *Journal of Library Administration*, *60*(2), 155–166. https://doi.org/10.1080/01930826.2019.1677092

Mackinder, Lisa. (n.d.) *Lisa Mackinder* [LinkedIn page]. LinkedIn. Retrieved April 17, 2023, from https://www.linkedin.com/in/lisamackinder/.

Newyear Ramirez, J. A., & Silva, J. (n.d.). *Serials reduction best practices and criteria* [Google Docs; UCB access]. Retrieved February 26, 2023, from https://docs.google.com/document/d/1AhMRWHLnl78XGi4V0JcbAu1WgvocQy9MI-migzJ0cyg

U.S. Department of Education, National Center for Education Statistics. (n.d.). *Integrated postsecondary education data system* [Data set]. https://nces.ed.gov/ipeds/

UC Berkeley Library Communications. (2022). *Moffitt Library is back; other UC Berkeley libraries have modified hours and services*. UC Berkeley Library. https://www.lib.berkeley.edu/about/news/2022-23-services

UC Berkeley Library. (n.d.). *Borrow + renew*. Retrieved February 5, 2023, from https://www.lib.berkeley.edu/find/borrow-renew

UC Berkeley Library. (2021). *Annual report 2020-21*. https://www.lib.berkeley.edu/sites/default/files/uc_berkeley_library_annual_report_2020-21.pdf

UC Berkeley Library. (2022). *Library facts and figures*. https://www.lib.berkeley.edu/sites/default/files/files/UCBLibraryFacts.pdf

UC Berkeley. (n.d.). *History & discoveries*. Retrieved February 5, 2023, from https://www.berkeley.edu/about/history-discoveries/

UC Libraries. (n.d.). *Systemwide ILS (SILS)*. UC Libraries. https://libraries.universityofcalifornia.edu/sils/

UCB Scholarly Resources. (2018). *2018 Reduction*. Library guides at UC Berkeley. Retrieved March 24, 2023, from https://guides.lib.berkeley.edu/c.php?g=706910&p=7986099

UCB Scholarly Resources. (2020). *Overview of the UC Berkeley Library collections budget*. Library Guides at UC Berkeley. Retrieved March 24, 2023, from https://guides.lib.berkeley.edu/scholarly-resources/reduction2020-budget

UCB Scholarly Resources. (2022). *2022-24 Budget reduction plan*. Library guides at UC Berkeley. Retrieved March 24, 2023, from https://guides.lib.berkeley.edu/scholarly-resources/2022_24_BudgetReduction

University Development and Alumni Relations. (2022). *CalFacts*. UC Berkeley. https://admissions.berkeley.edu/wp-content/uploads/K15224-Cal-Facts-2022_small.pdf

University of California. (2022, February 10). *Fall enrollment at a glance*. https://www.universityofcalifornia.edu/about-us/information-center/fall-enrollment-glance

Table 1

Most popular titles by percentage of total downloads, March 1 through March 26, 2023.

Rank	Title	Downloads	% of Total
1	*Vogue*	97,907	9.95%
2	*Journal of the American Chemical Society*	49,102	4.99%
3	*Analytical Chemistry*	19,874	2.02%
4	*The Ladies' Home Journal*	17,219	1.75%
5	*Journal of Organic Chemistry*	16,261	1.65%
6	*Harvard Business Review*	14,964	1.52%
7	*Organic Letters*	13,325	1.35%
8	*Nature Catalysis*	13,212	1.34%
9	*Nano Letters*	10,776	1.09%
10	*Environmental Monitoring and Assessment*	9,178	0.93%

Note: Figures taken from Alma Analytics report provided by Systems and Discovery Services Librarian Jackie Gosselar (personal communication, March 27, 2023). Analytics are not Mx. Gosselar's specialty and it's only a guess that this is in fact the period covered by this report, but a back-of-the-envelope comparison of the cost per use entailed by that assumption with the cost per use of CDL systemwide licensed resources suggests the assumption is at least plausible: 984,205 downloads / 26 days = 13.8 million downloads per year, which against the $13.8 million serials budget reported by the Library to USDOE yields a cost of exactly $1/download, while CDL reported 43.9 million systemwide downloads in 2019, against Mackinder's (n.d.) reported budget "in excess of $45 million," for a cost of approximately $1.03/download.

Of course the CDL numbers are for different years and the Berkeley numbers don't account for academic-year usage patterns or for print serials, but we're probably in the ball park.