Keeping Current

Reviews and Analysis of Special Reports

This column provides in-depth analysis of recently issued white papers, research studies, presentations, and special reports that the editor feels would be of greatest interest to academic librarians. This is a highly selective compilation, and the editor welcomes suggestions from readers <**Leslie_Stebbins@post.harvard.edu**>.

"Supporting the Changing Research Practices of Historians," by Jennifer Rutner and Roger Schonfeld. Final Report from ITHAKA S+R. December 10, 2012. Available: http://www.sr.ithaka.org/research-publications/supporting-changing-research-practices-historians

Overview

The authors conducted interviews targeted at the changing research habits of academic historians. The focus was on improving library services to historians. The research was based on interviews with almost 40 historians and dozens of librarians, archivists, and other service providers. The study found that while the overall work of historians has not changed fundamentally, the day-to-day work of historians has shifted significantly. The report makes a number of recommendations focused on how libraries can better support the day-to-day work of historians.

Key Points

- Libraries need to improve access including developing better finding aids, increasing digitization and integration of discovery tools, and improving promotion and teaching related to source materials.
- Format preferences would be addressed as well as increasing collaboration efforts to maximize access to archival materials. Greater staff expertise in sub-disciplines within history would be optimal for providing meaningful consultative services.
- Database providers need to evaluate the role they should play in light of Google's services.
- Citation software providers need to expand services to historians that address their specific needs related to gaining intellectual control of sources and supporting meaningful organizational schemes.
- History departments need to provide improved training in proposal development especially related to resource constraints and in the use of non-textual sources and new forms of scholarly expression.
- Scholarly societies need to more actively track changing needs of researchers.
- Funders need to explore ways to support connections between historians and their research support providers.

Analysis

This report is the first in a series of reports from ITHAKA aimed at understanding the changing research methods and practices of academic researchers. Focusing on academic historians, the authors have found a field that has changed significantly by expanding in subject areas, broadening in international scope, and using new technologies to more effectively conduct historical research. While the analysis below touches on some of the findings, reading the report in total is strongly recommended in order to fully understand the changing work of historians and the impact this has on planning library services.

The focus of an historian's work rests on their use of primary sources, and unfortunately many primary sources that historians need have yet to be digitized. Those interviewed felt the archivist played an essential role in helping researchers access valuable resources, but historians now interact with these materials in dramatically different ways that impact the contact they have with archivists. Digitized finding aids, digitized collections, and digital cameras all play important roles in the day to day work of carrying out historical research. The work that goes on in an archive today often focuses more on gathering and photographing and less on the analytical work that takes place at a later time outside of the archive. While providing the historian with some significant advantages, archives have not always recognized this change in work habits, and this separation of photographing work from analytic work impacts the services now required by historians.

Going physically to an archive is still a common part of most historical research projects, but funding and travel time are typically a challenge. The availability of digitized finding aids and digitized collections has greatly aided historians in their work. Having a good finding aid can sometimes eliminate the need for multiple trips to an archive, and even help researchers rule out a trip that would prove fruitless.

Most historians look for finding aids via Google, and it is clear that more digitized finding aids would greatly aid the work of historians. Locating primary sources is still a significant challenge for researchers, whereas the process of finding secondary literature has been greatly enhanced. Most historians are able to navigate the secondary literature handily and many find *Google Books* in particular extremely valuable. The subject knowledge of archivists about their own collections provides historians with invaluable support, whereas the subject knowledge of reference librarians is viewed as not specialized enough to be useful to historians. These researchers tended to value librarians for their teaching, interlibrary loan, and collection support roles rather than their research assistance, due to the specialized nature of most historical research.

The use of digital cameras and scanning devices seems to be the most important change in the practices of historians today. Capturing source material to be reviewed at a later time has greatly improved research efficiency and flexibility. Many historians had suggestions on ways archives could provide support for capturing source material. One of the largest challenges relates to the organizing of documents captured by digital cameras. These documents typically lack any

metadata and are often in JPEG format. In addition, there are no customized tools that easily help scholars combine their notes on a primary source with the images they are collecting while visiting an archive.

Using new technologies such as GIS, text mining, Google Ngram, visualization tools, and newer methods of scholarly communication are not yet widespread among historians, but they are areas that may be on the horizon for historical researchers. Collaboration among different staff members with technical expertise, as well as the need for more education and training, is viewed as important strategies to pursue as these new technologies gain traction.

"How Readers Discover Content in Scholarly Journals," by Tracy Gardner and Simon Inger. Renew Training. 2012. Available: http://www.renewtraining.com/How-Readers-Discover-Content-in-Scholarly-Journals-summary-edition.pdf

Overview

This report provides the results of an extensive survey of journal readers on the subject of journal content discovery. The survey of almost 20,000 readers around the world took place in the spring of 2012. Half the respondents came from the United States and United Kingdom, and about two-thirds of the respondents came from academic institutions. There was significant representation from all academic disciplines. This work builds on two previous surveys conducted in 2005 and 2008 by the same authors, and it analyzes trends over time in how readers discover content in scholarly journals. The report focuses on three types of reader behavior when searching for journals: citation searching, journal browsing, and subject searching.

Key Points

- The purpose of the report is to understand how readers use a journal website and inform publishers as to where readers begin their searching and how best to increase reader traffic.
- Bibliographic databases are the most popular general search engine used, followed closely by Google Scholar.
- Citation searching by researchers focused less on library web pages than it has in previous years and much more on going directly to a journal's homepage or an important research group or scholarly society web page. Alert services are still used heavily by researchers wanting to stay up to date on the latest articles, but these services have declined in the past seven years. At the same time there has been an increase in readers receiving alerts directly from a journal site.
- When subject searching for articles, specialized article databases remained by far the most popular method, followed by using a web search engine. Library web pages have increased in popularity as a method for subject searching and the use of scholarly society web pages are also increasing.

Analysis

One of the behaviors investigated by this extensive study involved asking what a researcher was doing when searching for the last article they pursued. A large majority of respondents were subject searching. The next two significant responses were either pursuing a journal alert they had already received or tracing a citation from another source. Interestingly, social media and email links were rarely mentioned by the respondents, though researchers from Asian countries tended to follow links in emails more frequently than researchers in North America and Europe. The survey also found that community websites such as Mendeley and ResearchGate have yet to catch on with most academic researchers.

While this survey was for the benefit of publishers engaged in redesigning their websites, librarians can greatly benefit from the

findings. The role played by library websites in discovering journal articles is still in a period of significant transition. A great deal of time and money has gone into managing electronic resources and improving access to ejournals via library web pages. With the advent of more and better web-scale discovery services,[1] libraries are embarking on an exciting journey that has the potential to provide valuable services to a research community in need of the appealing combination of massive searching within a quality filtered collaborative collection of resources. There is still a great deal to sort out and refine within these new web-scale discovery systems—issues such as transparency, vendor neutrality, depth of indexing, and stronger collaborative efforts on the part of libraries—but as these challenges are surmounted the results of the next survey of reader behaviors could show a dramatically altered landscape in terms of the role played by the library versus the journal publisher web sites.

"Manipulating Google Scholar Citations and Google Scholar Metrics: Simple, Easy and Tempting," by Emilio Delgado López-Cózar, Nicolás Robinson-García, and Daniel Torres Salinas. 2012. EC3 Working Papers 6: 29 May, 2012. Available: http://arxiv.org/abs/1212.0638

Overview

Google Scholar provides citations and metrics that provide researchers with easy access to bibliometric measuring. With this easy access comes the ability to manipulate the data and impact a scholar's H-index. The authors created a fake research group and six documents that cited the research group and uploaded these to a personal website under the University of Granada's domain name. The experiment resulted in an increase of 774 citations in 129 papers and increased the H-index for these authors and journals. The authors analyze the problems inherent in Google Scholar Metrics and Google Scholar Citations. (The H-index provides a measurement of the productivity and impact of a researcher and is based on a calculation involving the scholar's most cited papers and the number of citations they have received in other publications. The H-index can also be used to evaluate journals. It has meaning within specific fields of study but not across fields because citation practices vary widely.)

Key Points

- Google Scholar was launched in 2004 followed by Google Scholar Metrics that ranks journal impact and Google Scholar Citations that measures scholar output and impact. They are direct competitors to Thomas Reuters' Web of Science and Elsevier Scopus. They differ in that they work in an open environment pulling in materials from unfiltered repositories that are not screened for content. This leaves the door much more widely open for potential abuse and manipulation.
- Because self-citation and the manipulation of citation data are common, the authors analyzed Google Scholar's capacity to detect the manipulation of data.
- Unlike previous research on citation manipulation, the authors uploaded fake documents to their personal website that cited the documents of a research group. They did this in the simplest way possible by cutting and pasting repetitive text into the documents. The study also investigated Google Scholar's ability to detect retracted documents and delete their bibliographic records and list of references.

---

[1] For a useful update on web-scale discovery services and the implications for libraries see: "Library Web-Scale," by Marshall Breeding. *Computers in Libraries*, January 2012 and "Stakeholders Strive to Define Standards for Web-Scale Discovery Systems," by Michael Kelley. *Library Journal*, October 11, 2012.

Analysis

Creating false documents to discover defects to bibliographic tools has been going on for many decades. What was clear from this recent research is that it is extremely simple to manipulate author and journal rankings in Google Scholar and its related tools. These authors were able to easily manipulate the system and increase the rankings for the fake authors and journal metrics. Interestingly, it took almost a month for Google Scholar to pick up and index their fake articles, probably because they were posted to an institutional website rather than a formal archive. Also, when the researchers removed the false research papers, Google Scholar did not remove the bibliographic data. The faulty links continued to link to a cached version of the papers long after they were removed.

Google Scholar has become an extremely attractive tool for researchers because it is free, current, comprehensive, and simple to use. The authors have argued in previous papers that Google Scholar needs to play a stronger role in minimizing biases and manipulative practices. But comparing Google Scholar to more traditional databases of indexed journal articles is a little like comparing apples to oranges. Vendor databases such as Web of Science select published journal articles to be indexed. This greatly filters the content that is contained in the database. Google Scholar aims for comprehensiveness without regard to providing a qualitative filter. In addition, vendor databases provide a policing role and will exclude fraudulent or questionable works. Google Scholar does not actively look for or investigate citation errors or fraud.

The research is useful in pointing out that with very little effort the H-Index for authors and journals can be easily manipulated in Google Scholar. But the key takeaway is not necessarily the expectation that Google Scholar needs to address this issue. It would do so at the expense of being comprehensive, which is what makes it useful. The takeaway is that at present Google Scholar is not a good resource for measuring impact factors. Google Scholar is no more in a position to be able to accurately evaluate research output and impact in the same way that placing a paper in an institutional archive provides no guarantee of quality or peer review. Google Scholar and its related citation tools are too easily gamed to be useful for reliable qualitative bibliometric data.

"Academic Libraries and Research Data Services: Current Practices and Plans for the Future," by Carol Tenopir, Ben Birch and Suzie Allard. 2012. Chicago: Association of College and Research Libraries. Available: http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/Tenopir_Birch_Allard.pdf

Overview

The purpose of this report is to provide an analysis of the current and future plans for research data services in academic libraries. The report provides a strong argument for the implementation of more robust research data services in institutions of higher education and points to the central support role academic librarians can play. The study was comprised of a survey of ACRL library members in the United States and Canada to provide a current snapshot and future plans of library research data services in the sciences. The report also provides broad recommendations for library directors planning to expand or implement research data services.

Key Points

- Few academic libraries offer much in the way of research data services, but close to one-third plan to offer some services within the next two years.
- The most common service provided by libraries currently is creating web guides to help researchers locate data.

- Larger universities are more likely to offer consultative services and be involved in the knowledge creation process. Also, libraries on campuses that receive NSF funds are more apt to offer research data services. This suggests that it is likely that funding agency requirements are helping push the expansion of more robust research data services.
- For libraries to be successfully involved in research data services it makes sense for them to embark on collaborative efforts with other units on campus. The most likely partner for an academic library is the campus office of research.
- Reassigning work to existing staff is the most frequent method used for providing research data services. It is rare, though not impossible, to fund new positions for these services. Libraries are relying on workshops and conferences to provide professional development opportunities for assisting staff in learning skills needed to support data services. The report urges those farther along in the process, in addition to professional associations, to help other libraries get up to speed.

Analysis

This report provides a useful picture of the current state of research data services offered by libraries and also provides suggestions for actions library administrators can take to create and expand these services. The suggested actions are somewhat generic and broad, but would provide a useful outline for any library embarking on an implementation plan or expansion of services. The suggested actions include a careful scan of the e-science environment on campus, deciding whether to create new services or expand traditional ones, examining data services for the sciences that can be expanded into other disciplines, pursuing professional development opportunities for staff, and collaborating with other offices on campus to help position the library as a key player.

While funding agencies in the sciences are now requiring better data management, there is a gap between what is needed and the skill level and background of faculty, graduate students, and librarians. While the library is a natural choice for the organization and management of information—from traditional books and electronic resources to data sets—librarians can only provide meaningful support and expertise if they are able to acquire greatly needed professional development in the areas of direct support for faculty, consultative services, and infrastructure development. Librarians already have valuable skills related to collaborating with departments and offices, working across disciplines, and providing customer-focused services.

Essential to moving forward is the need for more robust professional development opportunities for librarians. For larger institutions especially, there is a need to provide librarians with more than a handful of workshops. The libraries at Worcester Polytechnic Institute and University of Massachusetts Medical School recently developed a full modular curriculum that takes a case-based approach to teaching data management.[2] This curriculum seems like a natural for getting funding and turning into a MOOC. With one-third of libraries planning to institute significant data management services in the next two years, there is a strong need for substantive professional development programs.

Leslie Stebbins
E-mail address: Leslie_Stebbins@post.harvard.edu.

10 January 2013

[2] See: Frameworks for a Data Management Curriculum at http://library.umassmed.edu/data_management_frameworks.pdf.