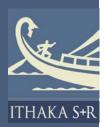REPORT

Barriers to Discovery of and Access to the Scientific Literature in the Corporate Sector

June 16, 2016

Roger C. Schonfeld

ITHAKA S+R

Ithaka S+R is a strategic consulting and research service provided by ITHAKA, a not-for-profit organization dedicated to helping the academic community use digital technologies to preserve the scholarly record and to advance research and teaching in sustainable ways. Ithaka S+R focuses on the transformation of scholarship and teaching in an online environment, with the goal of identifying the critical issues facing our community and acting as a catalyst for change. JSTOR, a research and learning platform, and Portico, a digital preservation service, are also part of ITHAKA.

Introduction

The offices that provide scientific information resources in major research companies – sometimes still called libraries – seek to make these resources available as seamlessly as possible. Barriers exist both for discovery and access. Corporate information professionals also face barriers in administering these information resources. In this project, Ithaka S+R examined the degree to which these various barriers are seen to exist in pharmaceutical, chemical, and food companies. The project had as its objective to inform the design of solutions from publishers, intermediaries, and scientific research companies themselves, to improve the discovery of and access to the scientific literature in corporate settings.

Ithaka S+R has been examining how researchers discover and access scholarly literature, and the barriers they face in doing so, for more than a decade. We have been especially interested in how user behaviors and work practices (which we sometimes call "workflows") should shape, and can be influenced by, the strategic directions of libraries, publishers, and intermediaries. We have followed the growth of Google and Google Scholar for research discovery, the development of library discovery services, the growing need for offsite and mobile access, and the challenges that even open access materials face in various types of workflows. Publishers want to maximize the usage of their content to establish their value proposition and drive revenue in a changing marketplace. Libraries want to serve the full information needs of their users while controlling costs. Ithaka S+R's contributions have helped to bring focus to the overall practices and workflows of researchers, where real barriers and limitations can be found.[1]

But for many publishers, especially STM publishers, higher education is only one marketplace in which they operate. Their journals, books, and other content may also be sold to healthcare organizations and within various corporate sectors. Research objectives and practices, organizational policy, and regulatory considerations vary substantially here compared with the academic sector, and research workflows, collaboration needs, and library service vary accordingly.

---

[1] See for example, Roger C. Schonfeld, "Meeting Researchers Where They Start: Streamlining Access to Scholarly Resources," *Ithaka S+R Issue Briefs*, March 26, 2015, available at http://dx.doi.org/10.18665/sr.241038, and Roger C. Schonfeld, "Does Discovery Still Happen in the Library? Roles and Strategies for a Shifting Reality," *Ithaka S+R Issue Briefs*, September 24, 2014, available at http://dx.doi.org/10.18665/sr.241038. An Ithaka S+R researcher previously examined corporate libraries from a different perspective. See Ross Housewright, "Themes of Change in Corporate Libraries: Considerations for Academic Libraries," *portal: Libraries and the Academy* 9, no. 2 (2009): 253-271, 10.1353/pla.0.0048.

With this background in mind, Ithaka S+R took on this project to examine how researchers within the corporate sector discover and access scholarly literature. We surveyed and interviewed librarians and other corporate information managers from nearly 40 companies, to understand where they perceive barriers exist. We also sought their opinions on the types of organizational and structural issues that would need to be considered in order to address these barriers. Focusing on researchers within the pharmaceutical, agriculture, and chemical sectors, we found that those in smaller organizations faced different challenges than those in more mature midsize and multi-national corporations.

Project structure and acknowledgments

This project was initiated by Copyright Clearance Center, which asked that Ithaka S+R consider expanding its research to cover the corporate use of scientific and scholarly information. We structured this project as sponsored research, with CCC serving as sponsor of the project and also joining the advisory committee. We gratefully acknowledge CCC's sponsorship of this project and also note that Ithaka S+R retained both complete control of the research and editorial independence for the final report.

An advisory committee reviewed preliminary plans for the project, provided input on data-gathering instruments, and reviewed a draft of the final report. Its members were:

- Tracey Armstrong, CCC

- Andrew Clark, UCB

- Helen Malone, GSK

- Babis Marmanis, CCC

- Deanna Marcum, Ithaka S+R

- Lauren Tulloch, CCC

- Sven-Olaf Vogt, Roche

Ithaka S+R's Kimberly Lutz reviewed a draft of this report and Liza Pagano assisted with the design of the figures.

Methodology and descriptive statistics

The population for this project include librarians and other corporate information professionals at pharmaceutical, food, and chemical companies. We sampled all the organizations represented in the membership of the Pharmaceutical Documentation Ring, as well as individuals from other companies identified by the members of our advisory committee.[2]

A survey questionnaire was developed and distributed to 122 librarians and other corporate information professionals in March 2016. The questionnaire is included as Appendix A.

Several reminder messages were sent to non-respondents. We received a total of 39 responses, of which three indicated they were the wrong contact individual and so the invitation was reissued to a colleague. In total, 36 usable responses were received, for a response rate of approximately 30%. Notwithstanding this healthy rate of response, the survey sample was small, and so the responses are analyzed with care as part of a mixed-methods approach further described below.

Several questions were included to allow for the stratification of survey findings. One that proved useful for many questions was an item asking "Approximately how many researchers and other individuals do you serve who are likely to be using scientific literature?" In analyzing the survey findings in this report, a stratification is provided by those respondents whose companies employ fewer than 1,000 such individuals and those who employ more than 1,000 such individuals. While it is not a perfect mapping, this seems to approximate the difference between small organizations and that have more of a start-up dynamic on the one hand and those that are midsize or large multinational organizations on the other. For example, every company with more than 1,000 such employees reported having them located in more than 20 countries, while roughly two-thirds of those with fewer than 1,000 such employees were in fewer than 20 countries. Of the 27 respondents that identified the number of researchers, there were 17 from companies with fewer than 1,000 of these employees, and 10 respondents from companies with more than 1,000. Because the smaller companies were well-represented in the survey but only six of their respondents expressed a willingness to be interviewed, a greater share of the interviews were conducted with individuals from larger companies.

---

[2] "The P-D-R is an association whose members represent the scientific information departments of the leading international R&D-based pharmaceutical corporations." For more information, see http://www.p-d-r.com/content/

Interviews of approximately 30-60 minutes were conducted in April 2016 by telephone. Interviewees were invited from survey respondents who indicated they would be willing to participate in a follow up interview, and ten such interviews were conducted in total. They were semi-structured, allowing interview subjects to steer the conversation in the directions that were of greatest relevance to their organizations. The interview protocol is provided as Appendix B and a list of interviewees is provided as Appendix C. The researcher took notes but did not record the interviews.

Organizational dynamics

Librarians and other corporate information managers saw several common considerations across all organization types. They are mostly focused on supporting corporate R&D functions, which at this sample of companies is scientific in nature. They are only sometimes, and then secondarily, charged with supporting other corporate activities, from marketing and sales to general counsel.

The focus of these libraries and corporate information services was on timely access to relevant and appropriate information resources at the lowest possible expense. In the cases where the interviewees are building library collections, principally of licensed scholarly journals and a variety of other databases and tools, the collections are intended to facilitate access rather than as being assets in themselves.

Beyond these common factors, there emerged two different sets of challenges that the companies face. Given the nature of the project's sample, these appear to align by whether the company is a comparatively smaller firm or an established multinational. This categorization is followed in the next two sections, exploring the characteristics of these two types of companies and some key differences between them.

Smaller companies

The small companies examined in this project are typically focusing on only one product, or a very small number of interrelated products. As a result, their researchers, and indeed other employees, tend to be focused on similar scientific and medical issues. They are growing organically and their information challenges are also growing organically. Solutions do not develop organically however but require punctuated intervention.

The small companies are too small to contemplate dedicated information management staffing. Someone on the research or medical affairs team—sometimes a senior leader with a PhD or an MD—may take responsibility for information management, as a small part of their job. They typically recognize the limitations in their ability to serve in this

role effectively and find it a distraction, even if a necessary one, from their core responsibilities.

In these settings, article access and information management are the most substantial challenges.

Article access is a challenge for these companies, the smaller of which subscribe to few journals and other information resources. One small organization that sees itself as a start-up reported that it subscribes to less than two dozen individual journals, a few general science titles and some of the core journals in the specific subfields where its research is concentrated. In such cases, the great majority of access comes through document delivery services and article purchase from publisher sites. Article rental models appeared to be a far less significant source for access.

Information management is a challenge for related reasons. The smaller companies are relying on article level access for a substantial portion of content needs, rather than licensed resources that can remain available over the course of time. Consequently, they need to manage the articles they acquire. Many an article would be useful to many other scientists, not just the individual who found it, because so many of the scientists are working in closely interrelated areas on a single product or a small number of them. How can they make these articles discoverable and accessible throughout the company? What are the copyright, fair use and intellectual property ramifications? The smaller companies in the sample reported using a variety of tools within their own organizations; at the smallest scale, a single individual ran a personal EndNote library and responded to requests on behalf of other researchers across the company, while there were also a variety of somewhat more scaled-up solutions.

Multinational corporations

Multinational companies have a global footprint, with multiple sites for research, and a marketing and sales presence on many continents. In addition to the overall challenges of all the multinational companies in managing a distributed scientific research apparatus, pharmaceutical companies face the added complexity of substantial regulatory commitments in a diversity of worldwide jurisdictions.

Library or corporate information management functions are often organized into some kind of R&D support organization. This organizational placement serves to emphasize the primacy of their role in support of research, although they may also maintain responsibility for the information needs of other parts of the company. Their overall mandate or mission includes supporting researchers' information needs, finding efficiencies that allow researchers to focus on their primary work, and building

connections across researchers that might result in new discoveries and ultimately new products. Their sister units often focus on topics such as corporate knowledge management, data management, or laboratory notebook provision.

Library or corporate information management units are focused exclusively on providing access to the scientific literature, while in other cases they also are connected with other units that provide specialized research functions. These include professional search services, literature reviews, and so forth.

These information management units are almost always smaller than they once were in terms of staffing, typically composed of one to 10 employees, with little apparent relationship to the size of the company. They rarely retain a physical collection or reading room, although in at least one case such a space has been retained. They may spend millions of dollars, as much as $10-20 million annually, on licensed e-resources.

Budgetary issues are paramount in these multinational corporate libraries. The interviews produced an array of discussions about how overhead and chargeback models influence the service and product needs of these libraries. For example, document delivery is often charged back to the unit whose researcher needed the document, putting the library in the position of facilitator rather than budgetary steward for content. In some cases, certain licensed e-resources are allocated to units for chargebacks based on usage metrics, which is a key factor driving interest in gathering more granular usage data, as discussed below.

Because article purchase and document delivery are seen as viable alternatives to licensing journal collections, many of the interviewees conduct periodic reviews to determine whether per-article models or licensed models will bring them the lowest cost. They value tools that help them run such analyses on a somewhat automated basis.[3]

Information security seems to be a greater concern at multinationals than at startups. Several interviewees mentioned that they have bilateral agreements in place with at least some content vendors to ensure that data about their usage of the scientific literature are not used by their vendors in a way that competes with them. One interviewee mentioned that her library is prohibited from uploading the holdings and IP address information to Google Scholar because they view even information about their network range and what they license as proprietary and not to be shared with Google. This in effect closes off

---

[3] On the other hand, at some large companies, there is no budgetary flexibility for mid-year content licensing, since budgets are allocated back to operating units as overhead. In these cases, demand for new resources is typically picked up by another unit at least for the first year.

some of the potential value in Google Scholar as a discovery and access tool for employees.

Finally, in terms of information security, one interviewee mentioned that she was forced to input login credential 10 or more times during the course of the day, to use a variety of business systems. For her, additional logins during a research workflow did not seem of great concern.

Several multinational interviewees emphasized that IT resources and support in their companies are outsourced, which produces an array of challenges. In such cases, IT provides only those specific functions that are included in the outsourcing agreement. As a result, it is far more difficult, if not impossible, for the library to develop an array of new services and internal partnerships for research support.

Workflows

The workflows that researchers experience vary across companies but the variance is especially prominent in comparing large multinational companies with smaller organizations.

Figure 1 illustrates a typical larger company researcher workflow, focused on tools provided by the library or corporate information center. These companies are building research libraries in their corporate areas of specialization.

One important starting point for researchers, simply labelled "database" in the illustration, can include any of a number of different tools and services:

- Business Source Corporate

- Embase

- Ovid

- Knovel

- PharmaPendium

- Reaxys

- SciFinder

- Scopus

- Web of Science

Library discovery and traditional link resolver services included those offered by ProQuest (including through its purchase of Ex Libris), Ebsco, and TDNet. Interviewees used document delivery services, sometimes with integrated link resolvers, from CCC's RightFind and ReprintsDesk's Article Galaxy, built into the workflow to the greatest extent possible.
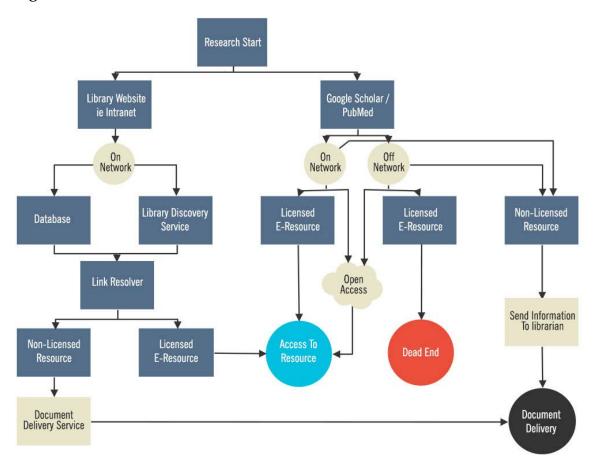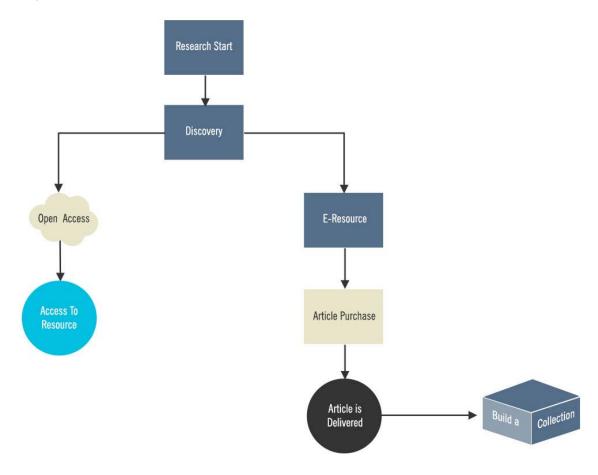
Figure 1

Figure 2 illustrates a typical researcher workflow at a smaller company, which is almost entirely driven by self-service article purchase. In one model, they purchase articles and build up a collection of purchased articles using EndNote, SharePoint, or similar tools. They look forward to being able to build a more integrated workflow including services that integrate document delivery, collaboration, and content management, such as Article Galaxy, Bibliogo, and RightFind. Small companies may typically subscribe to a small set of journals, a combination of general science titles and field-specific titles, which are especially important for current awareness.

Figure 2

The following sections are organized from information discovery to article access, focusing in on some of the key issues that arose in the survey and interviews.
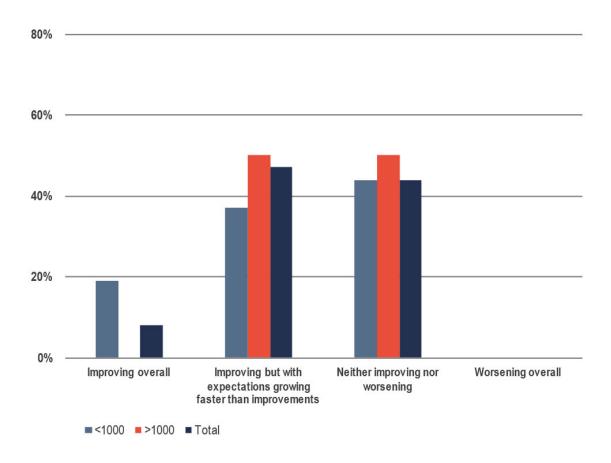
Starting points impossible to control

Every librarian and other corporate information professional interviewed expressed the conviction that they had to be able to support essentially all possible starting points for their researchers. At small companies, there was no sense of interest in or concern with controlling starting points. At larger companies, information professionals recognized that there were benefits to users if they started from the library homepage, because, as illustrated in Figure 1, that would take users through the access authorization process seamlessly. While in several cases, the library made an effort to urge or steer researchers through this site, there was a recognition that a diversity of other starting points are being utilized as well.

Discovery seen as static or modestly improving

Survey respondents characterize scientific article discovery at their organization as neither improving nor worsening overall, but rather static or modestly improving. As Figure 3 indicates, this pattern appears to hold across smaller and larger companies.

Figure 3

Various philosophies on information discovery

Interviewees expressed several discrete philosophies about the needs of their users in terms of information discovery. No obvious pattern emerged as to why some interviewees held one view or the other.
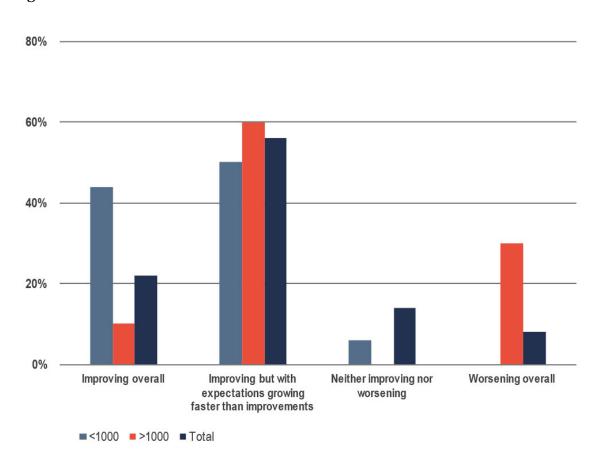
One set of interviewees pointed strongly to the need to have a single discovery service that included all possible scientific resources. They are investigating or have acquired the single-index discovery services such as those provided by Ebsco, ProQuest, and TDNet. The vision is a discovery starting point that indexes "everything." Various interview subjects suggested that its scope would need to include the complete scientific literature, which according to some is inadequately represented, and is especially complicated given that not all resources are licensed for the entire company in the cases of larger organizations. It would also need to include internal materials too, such as research reports, which are not made available publicly. This group of interviewees seems to imagine that the solution to its needs is the creation of a single starting point through which it could ultimately route all its researchers, addressing some of the downstream challenges discussed below.

Another set of librarians and corporate information professionals pointed to the inadequacy of a single discovery interface. These interviewees expressed confidence that different databases offer – and will continue to offer – unique discovery features that were highly valued by researchers. They pointed to controlled vocabularies specific to individual fields and specific interface features that are not broadly extensible. This group of interviewees seems to acknowledge that there will never be a single index or interface that can serve as a starting point, albeit the library website can serve as a starting point for accessing all these databases.

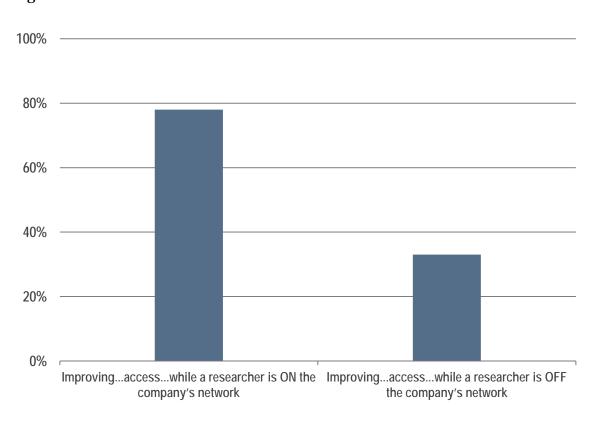*Expectations for access to licensed e-resources growing faster than improvements*

As Figure 4 illustrates, the highest share of respondents would characterize expectations for access to scientific articles as growing faster than the improvements they are seeing. Survey respondents from larger corporations are more likely to see access issues actually worsening. Those from smaller companies are more likely to see an improvement overall, which may reflect in part the successes in integrating document delivery into research workflows in recent years.

Figure 4

As illustrated in Figure 5, the survey found that a notably higher share of respondents felt improving onsite access was very important as compared with offsite access.

Figure 5

Onsite access: Mixed reviews for IP authentication

IP authentication plays a substantial role for many companies and some of its principal challenges are for on-site usage where administering these authentication systems can be a challenge.

At larger companies, corporate acquisitions and unit divestments are a regular occurrence with which the library must grapple, and in some cases have substantial implications for licensed e-resources. Depending on the specific nature of the transaction, individuals may be using a corporate IP address range, or have an account on the ActiveDirectory, even after they are no longer employees, as divestment processes can require a transitional period. Conversely, it may take time to set up the employees of an acquired company on the IP range or in the ActiveDirectory, resulting in access

management challenges. Some interviewees reported extensive workarounds needed to account for such dynamics.

On the other hand, some interviewees reported that challenges with IP authentication were not as serious as these divestments and acquisitions might suggest. Proxy servers are used by many corporations as part of their effort to ensure network security. They route all outbound and inbound internet traffic through a single or small number of IP addresses, masking which individual users might be associated with that information. Several interviewees indicated that these proxy servers are extremely valuable for another reason, because they only report to a content provider the IP addresses of these proxy servers, which rarely if ever change. As IP ranges are added or subtracted from the corporate network, for technical or organizational reasons, these are handled simply by adjusting the configuration of the corporate proxy. In an effort to lock down the network, the companies adopting this practice also make it far easier to administer IP authentication.

Some companies are trying to move beyond IP authentication to deliver true single-sign-on to content platforms. Interviewees that reported these efforts face the need to develop specialized work practices with each of dozens of content providers. The patchwork of authentication options that results may be less effective than remaining entirely with IP authentication.

Offsite access: Needs vary substantially

There was tremendous variation among interviewees in describing the level of need for offsite access to the scientific literature.
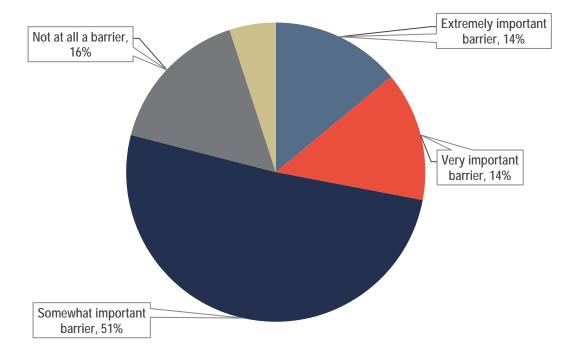
Some described a robust culture of working from home and/or a variety of other needs for offsite access. Of these, some felt that offsite access was adequate, because Citrix or VPN solutions were seen as being essentially satisfactory for the purposes of accessing licensed e-resources. Others felt that only a true single sign on solution could work for their needs.

Another set of interviewees felt that offsite use cases were unusual and idiosyncratic. These interviewees did not generally expect substantial growth in the need for offsite access because their organizational cultures are such that work happens at work. While sales staff and other functions may entail substantial travel, no great need was seen to provide off-site access to the scientific literature.

Mobile devices not seen as highest priority

While needs varied for offsite access, mobile devices were less frequently reported as a substantial concern. Less than a third of survey respondents depicted "Limited ability to support mobile and non-PC usage" as a very or extremely important barrier, as illustrated in Figure 6, substantially lower than some other barriers. Interviewees were generally in agreement. One interviewee said in surprise that he had never heard of anyone trying to read a scientific article on a mobile phone. Of course, there are other workflow use cases, such as discovery, for small-form devices.

Figure 6

Access: Document delivery vital but sometimes a bottleneck

Many companies rely on document delivery as an interchangeable feature with subscription access, used whenever licensing fee relative to usage does not justify a subscription. New services have in recent years offered stronger alternatives for document delivery that fit into established discovery and access workflows, using link resolver technology to offer document delivery options instead of denying access. Interviewees whose companies offer such services were generally positive about them, focusing in on the specific aspects of how effectively the workflow integrations were

established. Given the myriad possible starting points, it can be a challenge to effectively integrate link resolver-driven service offerings into every possible research workflow.

Interviewees accepted that document delivery will not always be instantaneous, since it sometimes requires some type of manual intervention. They nevertheless want these features to work seamlessly, and several emphasized that this step can be "messy." They want to see link resolvers, including those built in to document delivery services, provide complete coverage of the options at the point of publication and/or first online availability. Several interviewees emphasized that there is a lag in information reaching a link resolver, which can result in poor decisions being made about how to deliver an article optimally to a researcher.

Access: Price and open access

Issues connected to the price of licensing e-resources came up naturally on several occasions. One interviewee from a large multinational corporation said with complete confidence that they spent more on licensed e-resources annually than the largest research university. Budgets are managed very strictly at many of the large companies, and there is a keen awareness in many cases of the tradeoffs between licensing and document delivery approaches to providing content.

One issue that came up during several interviews was how the corporation is bounded in the license agreement. Some of the largest multi-nationals outsource a variety of functions that require extensive access to the scientific literature, including certain forms of regulatory submissions. These interviewees believe that whether work is outsourced or insourced should be immaterial to the total price of licensed access to the scientific literature. In some cases, these companies believe that, between what they pay directly for their employees and indirectly for such vendors, the cost of accessing scientific literature is too high.

Several interviewees brought up open access as a promising alternative to some of the access bottlenecks or pricing challenges they feel they currently face. They emphasized that a growing shift towards open access would help them to manage their companies' information needs.

At the companies studied in this project, there is typically a bright line separating library services, which are about content acquisition, from medical or scientific writing/publishing offices, which support external content dissemination. Even in smaller companies, where the same person may have ultimate responsibility for both content acquisitions and publishing, encouraging scientists to publish in open access channels has not been a priority. Interviewees believe that scientists value the reputation

of the journal in which they publish and that only a small set of titles have established a premier reputation while operating an open access business model.

While open access articles were cited by most interviewees as easing some of the access challenges, they generally did little to encourage their production.

Information management is a silent challenge

In interview after interview designed to focus on discovery and access, information management arose as a priority in one sense or another.

At smaller companies, the interviewees were challenged to manage on behalf of the company those articles that their scientists had purchased. In these cases, they posted articles to internal spaces such as a corporate intranet or used personal solutions including EndNote, to respond to email requests from colleagues. Services that would automate the functions of building an internal "collection" were seen as highly desirable but too complicated or expensive to operate internally.

At larger companies, the information management challenges were rather different. There, several interviewees expressed an interest in being able to do two things that are currently difficult or impossible to do. First, they wanted to identify productizable discoveries already in existence in the corporate archives (although the term archives was never used). Second, they wanted to connect scientists from across the global footprint of the company who were working on similar topics or might be able to collaborate usefully.

Unclear need for user-level usage data

Many interviewees discussed the importance of aggregated usage data, for two principal purposes.

First, usage metrics are utilized for decision-making about subscription strategies. While usage may be seen by some as a proxy for value, interviewees spoke principally about how such metrics could be used to calculate the efficiency of an unlimited-usage license versus relying on document delivery. This objective requires only aggregated usage data.

Second, in some corporate settings, usage data are a factor in establishing charge-backs to departments. This objective requires some mechanism to connect individual users, or appropriate groups of them, with their usage. While IP ranges sometimes permit a rough assignment of usage by department, network security such as proxies sometimes impeded the ability to collect such data. Several interviewees expressed a desire for usage

data on a more granular basis that would allow them to allocate costs more effectively. This is a factor driving some larger corporations to seek single sign-on and other mechanisms for generating these metrics.

But while there is a desire for these types of aggregated usage data, interviewees were generally mixed in terms of their actual desire to be able to gather and analyze user-level usage analytics. Some European interviewees discussed the regulatory need to protect the privacy of employees and said that it would be impossible for them to identify user-specific activities. Several interviewees mentioned that they like that SciFinder requires the use of individual logins because the library can see who is doing what within that tool. Others have aspirations to gather individual usage patterns across a broader array of content platforms, with the goal of providing better services and support, but they are unclear how exactly such data would be used or whether it could ever become their priority. There is currently little use of analytics at a user level, only in aggregated form for purposes of budget allocation and license price negotiation.

Reflections

Companies are not all the same. Unlike their peers at smaller companies, information professionals at larger companies are increasingly frustrated with access challenges.

Access is presenting a challenge not only in certain cases to users but also in terms of the ability of corporate information professionals to provide the metrics and chargebacks that are essential to their success. IP authentication is imperfect. It seems to work effectively when a corporate proxy server minimizes the administrative burdens of changing network IP ranges, but this in turn create challenges for tracking usage in a way that permits usage-based departmental charge-backs. Alternatives to IP authentication will need to be mindful of this very specific context, and to be successful will need to scale, avoiding the trap of requiring bilateral arrangements between content provider and customer.

Looking across the sector, offsite access may be a less significant challenge, but it still poses limitations in some corporations. Given broader changes in society, academia, and the workplace, it would make sense to anticipate that needs for offsite, non-PC, and multi-device access will only grow.

From a researcher perspective, access appears to be a challenge principally because there are so many discovery starting points that can be relevant. Substantial efforts are being made to integrate link resolvers and document delivery into the research workflow, to minimize the number of "dead ends" that one experiences. At the start of a research

session, if the user does not go through the library/intranet as starting point, the access experience can be degraded in a number of ways.

If there were a single discovery starting point, through which all researchers and other users could be routed, it might simplify some of these challenges. But while some companies are pursuing this objective, others do not believe it to be realistic. The need to integrate a variety of advanced discovery functionalities particular to different fields, and the need to search not only the published literature but also internal corporate sources, are two challenges to the vision of a single solution.

Offsite access and mobile devices arose as issues but they were not as widespread as might have been expected. Offsite access is highly important for some organizations but almost a non-issue at others. Providing access to the scientific literature in a form factor optimized for mobile devices is not widely seen as an immediate priority, although the project did not explore other uses of mobile devices such as discovery or and information management. Broader context may indicate that behaviors and needs will trend towards remote access and mobile devices.

The complexity of administering this environment, especially in the largest corporations with at most a dozen information professionals for tens of thousands of researchers, cannot be overstated. Administrative, budget, and other business issues are richly intertwined with the user experience that can be provided to researchers.

As in the academic sector, the ultimate objective is a seamless workflow for researchers. But the decentralized provision of content, the variety of available discovery services, and individualized researcher preferences and workflows, remain substantial challenges to such a vision.

Appendix A: Survey Questionnaire

Today, scientific literature is widely available online, but key barriers in the discovery of and access to this literature have also been reported. In this project, we are investigating the extent to which such barriers pose real challenges in the commercial sector. We are also examining how such barriers can be addressed and ultimately eliminated. We are exploring topics such as authentication mechanisms, mobile device usage, and services to support personalized discovery.

As a corporate information professional, you have unique insights into the needs for access to scholarly and peer-reviewed content across your organization as well as opportunities to improve this access. Thank you for participating in this survey.

Before we begin, we want to confirm that you are the appropriate person to complete this survey.

Are you familiar with how researchers at your organization are authorized and authenticated to gain access to subscribed content?
❍ Yes
❍ No

[If no, ask this question and end the survey]

Can you recommend the name and email address of a colleague who is familiar with this topic?

- Name

- Email Address

[If yes, the survey continues here.]

Looking across the provision of scientific research articles and other scientific information resources, how important are each of the following priorities at your organization.

Improving the discovery of scientific articles
○ Very Unimportant
○ Somewhat Unimportant
○ Neither Important nor Unimportant
○ Somewhat Important
○ Very important

Improving the ease of access to scientific articles while a researcher is ON the company's network
○ Very Unimportant
○ Somewhat Unimportant
○ Neither Important nor Unimportant
○ Somewhat Important
○ Very important

Improving the ease of access to scientific articles while a researcher is OFF the company's network
○ Very Unimportant
○ Somewhat Unimportant
○ Neither Important nor Unimportant
○ Somewhat Important
○ Very important

Improving the discovery and use of research data and other primary source information, such as libraries of protein structures and genetic code
○ Very Unimportant
○ Somewhat Unimportant
○ Neither Important nor Unimportant
○ Somewhat Important
○ Very important

Improving the effectiveness of evaluating, assessing, and otherwise working with the scientific literature (for example, interpreting confidence in the findings)
- ○ Very Unimportant
- ○ Somewhat Unimportant
- ○ Neither Important nor Unimportant
- ○ Somewhat Important
- ○ Very important

Improving access controls to handle changes in corporate structure such as acquisitions, mergers, and divestitures
- ○ Very Unimportant
- ○ Somewhat Unimportant
- ○ Neither Important nor Unimportant
- ○ Somewhat Important
- ○ Very important

Improving the ability of researchers to publish their own research findings in the scientific literature
- ○ Very Unimportant
- ○ Somewhat Unimportant
- ○ Neither Important nor Unimportant
- ○ Somewhat Important
- ○ Very important

Improving the management of research data generated by scientists
- ○ Very Unimportant
- ○ Somewhat Unimportant
- ○ Neither Important nor Unimportant
- ○ Somewhat Important
- ○ Very important

Improving the provision of personalized services for individual scientists and research priorities
○ Very Unimportant
○ Somewhat Unimportant
○ Neither Important nor Unimportant
○ Somewhat Important
○ Very important

If there are other vital priorities for improving the provision and management of scientific information, please share those briefly here:

For the following questions, please think specifically about the discovery of and access to research articles in the scientific literature.

How efficiently are scientists at your organization able to discover research articles?
○ Not at all efficiently
○ Somewhat efficiently
○ Very efficiently
○ Extremely efficiently
○ Don't Know/Not Sure

How comprehensively are scientists at your organization able to discover research articles?
○ Not at all comprehensively
○ Somewhat comprehensively
○ Very comprehensively
○ Extremely comprehensively
○ Don't Know/Not Sure

Having discovered a research article they would like to read, how readily are scientists at your organization able to access that research article?

❍ Not at all readily
❍ Somewhat readily
❍ Very readily
❍ Extremely readily
❍ Don't Know/Not Sure

To what extent do you believe that each of the following issues presents a barrier to the discovery of and access to research articles in your organization:

Authentication models
❍ Extremely important barrier
❍ Very important barrier
❍ Somewhat important barrier
❍ Not at all a barrier
❍ Don't Know/Not Sure

Limited number and breadth of resources that can be accessed through the same username and password
❍ Extremely important barrier
❍ Very important barrier
❍ Somewhat important barrier
❍ Not at all a barrier
❍ Don't Know/Not Sure

Off-site access limitations and challenges
❍ Extremely important barrier
❍ Very important barrier
❍ Somewhat important barrier
❍ Not at all a barrier
❍ Don't Know/Not Sure

Authorization models
- ❍ Extremely important barrier
- ❍ Very important barrier
- ❍ Somewhat important barrier
- ❍ Not at all a barrier
- ❍ Don't Know/Not Sure

Metadata limitations
- ❍ Extremely important barrier
- ❍ Very important barrier
- ❍ Somewhat important barrier
- ❍ Not at all a barrier
- ❍ Don't Know/Not Sure

Lack of information about usage patterns
- ❍ Extremely important barrier
- ❍ Very important barrier
- ❍ Somewhat important barrier
- ❍ Not at all a barrier
- ❍ Don't Know/Not Sure

Limited ability to support mobile and non-PC usage
- ❍ Extremely important barrier
- ❍ Very important barrier
- ❍ Somewhat important barrier
- ❍ Not at all a barrier
- ❍ Don't Know/Not Sure

Insufficient availability of reviews and syntheses
- ❍ Extremely important barrier
- ❍ Very important barrier
- ❍ Somewhat important barrier
- ❍ Not at all a barrier
- ❍ Don't Know/Not Sure

Preprints are used when a final version should be preferred
○ Extremely important barrier
○ Very important barrier
○ Somewhat important barrier
○ Not at all a barrier
○ Don't Know/Not Sure

Publisher and/or technology models are ill-suited to our organization's global footprint
○ Extremely important barrier
○ Very important barrier
○ Somewhat important barrier
○ Not at all a barrier
○ Don't Know/Not Sure

If there are other important barriers not listed above, please outline them briefly here:

Overall, how would you characterize discovery of scientific articles at your organization?
○ Improving overall
○ Improving but with expectations growing faster than improvements
○ Neither improving nor worsening
○ Worsening overall

Overall, how would you characterize access to scientific articles at your organization?
○ Improving overall
○ Improving but with expectations growing faster than improvements
○ Neither improving nor worsening
○ Worsening overall

When providing access to the research literature, do you utilize any of the following services:

❍ Shibboleth
❍ OpenAthens
❍ A proxy server
❍ A single sign-on provider
❍ Other federated identity management provider
❍ Other

Please indicate product name and provider information for your proxy server, single sign-on provider, other federated identity management provider, or other authentication/authorization service.

Looking across the publishers, content providers, and other vendors that you work with to provide scientific articles to your organization, which types of products and services work especially well?

Looking across the publishers, content providers, and other vendors that you work with to provide scientific articles to your organization, which types of products and services pose special barriers?

Please tell us a little bit about the organization for which you provide information services including scientific articles.

Approximately how many researchers and other individuals do you serve who are likely to be using scientific literature?

In how many countries are your researchers and other users of scientific literature located?

❍ 1
❍ 2-5
❍ 6-10
❍ 11-20
❍ 21-50
❍ 51-100
❍ More than 100

Please tell us a little bit about yourself. Which of the following best describes your role?

❍ Librarian
❍ Knowledge Manager
❍ Technologist
❍ Scientist
❍ Other

Beyond yourself, what other stakeholders at your organization are involved in these issues related to the discovery of and access to scientific articles? For example, IT, information security, etc.

Would you be willing to be contacted to participate in an interview to explore your answers and these issues in greater depth?

❍ Yes
❍ No

Appendix B: Interview Protocol

- Please describe the research environment at your organization [are researchers distributed or centralized; is their work principally in a wet lab or in other contexts as well; etc]

- What roles do you play and where do you fit into the organization

- Please describe some of the discovery processes that your researchers utilize for the scientific literature, including what works well and any limitations [prompt for issues related to privacy/awareness, library role in literature review]

- Please describe some of the research discovery processes that your researchers utilize for primary research data [prompt re need for seamlessness across internal and published sources]

- Please describe some of the information access processes that your researchers utilize, including what works well and any limitations [prompt for IP authentication; prompt for centralized solutions vs subgroup solutions vs. PPV models; prompt for link resolver to publisher packages vs other models; prompt for proxy servers and other technical means]

- To what extent do your researchers require access when outside the secure network, potentially using a non-company device?

- Is greater seamlessness an objective or is this overall working well?

- Please describe the usage data and analytics you receive. [prompt for desire to receive such data in greater granularity]

- What types of internal collaborations will be required in order to address these issues? [Prompt for IT/network security]

Appendix C: List of Interviewees

The following individuals were interviewed as part of the project:

- Stephan Ballenweg, BASF

- Rachel Benzies, Syngenta

- Tom Chittenden, ImmunoGen

- Liisa Eisenlohr, Clovis Oncology

- Jennifer Kooy, Abbott Laboratories

- Maria Markovic, Myriad Genetics

- Brian Schurrman, Novo Nordisk

- Peter Vogelman, Ferring Pharmaceuticals

Two additional interviewees preferred to remain anonymous.

The contributions of all interviews and survey respondents are acknowledged with gratitude.