Understanding the Research and Information Practices of Chemists at the University of Texas at Austin

Final Report

David Flaxbart
Liaison Librarian for Chemistry

Lydia Fletcher
Liaison Librarian for Physical Sciences and Mathematics

University of Texas Libraries

Fall 2018

SUMMARY

Key Observations

- Chemistry is highly interdisciplinary, with significant focus in life and health science applications as well as technology.
- Chemists are profoundly dependent on ready access to peer-reviewed journals as their primary source of information, as well as the outlet for their own publications. Use of other formats, including books, is low by comparison.
- While the volume of published literature grows relentlessly, faculty feel that its overall quality is declining, and it's becoming ever more difficult to find the gems amid the dross. Despite the proliferation of new journals, chemists – both as readers and authors – tend to focus on a relatively small number of "top" journals and publishers, particularly ACS and RSC.
- Keeping up with the flood of new papers is a problem. Faculty use a variety of coping mechanisms, but many expressed feelings of being overwhelmed and wondered if staying current is even possible.
- Management and later retrieval of articles downloaded in the course of literature work is frustrating and impeded by the lack of adequate tools.
- Open publication is understood as a generalized benefit, but it's not a priority for authors. Paying APCs to make articles open was dismissed as unnecessary or even foolish. Green OA is primarily an outcome of funding agency mandates and compliance, not of a proactive desire to be open. Chemistry lacks a preprint culture as well.
- Information-seeking strategies vary, within limits. They use a variety of search tools according to their tastes and experiences, but there is little innovation. Web of Science was the leading go-to discovery tool.
- Their needs for externally-sourced data (as opposed to text) are minimal, and limited largely to data published as supporting information to articles, and data from established disciplinary repositories.
- Data management strategies are ad-hoc, underdeveloped, or nonexistent.
- Faculty see little need to preserve or share raw/preliminary experimental data. Some datasets are easier to recreate at point of need than to store indefinitely just in case.
- Laboratory recordkeeping is almost entirely paper-based. There is awareness but little uptake of available support technologies such as electronic lab notebooks (ELNs), reference managers, data archiving tools, etc. There is not much incentive to innovate in these areas, and the forces of inertia, tradition, cost, and the lack of time are powerful obstacles to changing the status quo.
- There is little awareness of library support services or systems, such as local repositories.

Recommendations and Caveats

- Explore potential support services around discovery and management of new literature: alerts, reference/file managers, best practices, etc.
- Frame the open agenda in ways that emphasize positive outcomes rather than costs and consequences.

- The most impactful potential solutions to data management challenges are beyond the capability and resources of the library, and should be addressed by the campus and wider disciplinary community.
- Efforts to build awareness of data management best practices should be directed at individual lab groups that indicate explicit interest, with an emphasis on informing graduate students.
- When promoting or teaching more sophisticated discovery tools (e.g. SciFinder and Reaxys), focus on graduate students, who are more likely to try them and use them consistently.
- Proposals to support productivity "solutions" (such as ELNs) should be contingent on demonstrated demand and available resources.

INTRODUCTION

Supporting the needs of chemists has often been regarded as a special challenge for academic libraries. Chemists have long had a reputation, sometimes but not always deserved, of being a difficult population to reach and serve. They have high expectations of library-supplied resources, which are among the most expensive in any field. The unique nature of their specialized research tools demands a significant level of in-depth knowledge on the part of librarians tasked with serving their needs. They are also often resistant to the usual channels of engagement and outreach.

Chemists are no different from most academic scientists in having a singular focus on their own priorities. They are totally immersed in the things that matter most to them: pursuing individual lines of inquiry, acquiring funding to support that research, gathering a team of graduate students and post-docs to carry it out, publishing and disseminating their results, gaining tenure, and building an international reputation as an expert in their field. A second common trait is that they are very busy people. While many take their primary satisfaction from pursuing pure scientific inquiry, they also serve as teachers, mentors, administrators, and lab managers. They have little time left over to think about bigger-picture subjects such as scholarly communication, publishing trends, open science, etc.

Understanding the way chemists think and behave is the first step toward developing a suite of services and an engagement strategy that can be successful in the long term. Our hope is to use this understanding to identify specific areas of need where library intervention might be both fruitful and sustainable, as well as other areas where it may not be as effective. Libraries should strive to engage with chemists in their own milieu while avoiding library-centric messaging. Chemists, like all scientists, regard the library as a part of the supporting infrastructure for research, not as an end in itself.

The Chemistry Department

As of 2018 the Department of Chemistry, a unit of the College of Natural Sciences, had 28 tenured or tenure-track faculty members across four sub-disciplinary divisions: Organic, Inorganic, Physical, and Analytical chemistry. (Biochemistry split off from the department in 2013 to join the new combined department of Molecular Biosciences.) The offices and lab facilities of chemistry faculty are spread across several buildings but are centered in Welch Hall and the Norman Hackerman Building.

METHOD

This study follows on two prior chemistry-focused studies in particular: a smaller UT-Austin project carried out in 2002 (Flaxbart 2003) and an extensive Ithaka study of chemists in the United Kingdom, published in 2013 (Long & Schonfeld 2013). Both used an interview approach to assess information practices and attitudes in this academic population.

The present study adapted an interview-based methodology devised by Ithaka S+R, a not-for-profit research and consulting service that advises academic, cultural, and publishing communities. We developed a semi-structured interview script based on prior Ithaka studies, and submitted it for approval by UT Austin’s Institutional Review Board (IRB). (See the Appendix for the interview questions.)

We sought 15 tenure-track or tenured faculty members to interview, and to that end sent email invitations to 19 professors asking for their participation. Four of these did not respond. While the selection of potential participants was not random or scientific, it provided a good cross section of the department. All faculty ranks – assistant, associate, and full professors – across all four subdivisions were represented in the participating group. Interview subjects were assured of anonymity.

One-on-one interviews took place in faculty offices from February to April 2018. The interview sessions were recorded and later transcribed by the library's in-house transcription service. The interviews varied in length from 20 to 40 minutes. While following a predetermined set of questions, we were flexible with the way the conversations proceeded, asking follow up questions where appropriate and explaining specific terms and concepts where necessary, while allowing the interviewees to respond to the extent they felt comfortable.

The interview transcripts were lightly edited and corrected for clarity, and then analyzed separately by the authors. We sought to draw out common themes and concerns, make generalized observations, and highlight unique responses that encapsulated the attitudes and challenges faced by chemistry researchers today.

THEMES OF QUESTIONS

The interview questions were crafted to explore the research lives of chemistry faculty, from their own point of view rather than the library's. It was important to reassure them that the interviews would not be "about the library" per se – a topic about which most faculty are only minimally knowledgeable – but instead about their own work routines, practices and priorities. It is in their role as independent researchers and lab leaders (Principal Investigators, or PIs) that we approached them for this study; their other roles in teaching and wider service were not addressed in the interviews. The themes of questions can be summarized as follows:

- Background
    - Research focus/topics
    - Funding and collaborations
- Instrumentation and data output
- Lab recordkeeping
- Information Discovery and management
- Keeping up with trends and new research
- External data needs and discovery
- Publication practices
    - Open publication
    - Data sharing
- Research Data Management

RESULTS AND DISCUSSION

General Background

A. Research Focus

Each interviewee was asked at the outset to describe the topics of their current research. Most were adept at presenting a coherent and understandable summary without hesitation, indicating that they are practiced in this art of explaining to laypersons what they do. They described a wide variety of projects and broad areas of inquiry. Chemistry is often called "the central science" (there is even a journal by that name), so its interdisciplinary reach is extensive. Emphasis on direct application of research is now common, especially as related to life and health sciences, as well as in materials and energy-related technologies. Despite the administrative transfer of biochemistry faculty to Molecular Biosciences in 2013, research in the department's four remaining divisions is increasingly interdisciplinary, and has a significant biological component. There are widespread collaborations across departments and colleges, as well as with other institutions. While most chemists interviewed for this study carry out traditional wet-lab research, there are also theoretical chemists who approach problems via computer simulation and modeling.

Figure 1 is a visualization of "research areas" as defined by a Web of Science analysis of the recent publications of the 15 interviewees.

Figure 1. Heat map of research areas of Faculty interviewed.

Source: Web of Science

B. Funding

External funding for research projects comes from a relatively short list of "usual suspects" – predominantly U.S. federal agencies. Academic chemists in Texas also benefit from the generosity of the Robert A. Welch Foundation, a Houston-based philanthropic organization dedicated to supporting chemical research and education in the state. Many PIs in chemical sciences and technology at UT Austin have longstanding Welch grants. As expected, NIH, NSF, and Defense accounted for much of the federal funding described.

Table 1. Funding sources.

Funding Source	*No. of responses*
Welch Foundation	12
NIH	10
NSF	9
Defense	5
DOE	3
Industry	3
Other Foundations	5
Other	5

While few interviewees volunteered opinions about the state of research funding today, one chemist did complain that the scramble for grants is "a dark underbelly of academics [where] we spend all our time either working for or worrying about money." This is understandable, because grant funding is central to their career advancement.

C. Collaborations

Collaboration among researchers is widespread, and frequently crosses disciplinary and organizational lines. While much research in chemistry remains localized within the PI's lab group, it is rare that others outside the group are not involved in some fashion, including receiving co-author credit on papers. Interviewees were asked about collaborations on and off campus. Most faculty in STEM departments are affiliated with one or more cross-disciplinary "organized research units" that encourage and facilitate inter-departmental collaboration (examples include the Texas Materials Institute and the Institute for Cellular and Molecular Biology). On campus units named as sources of collaboration are summarized in the table below.

Table 2. Departments named as sources of on-campus collaborations.

Named Campus Collaborations	*No. of responses*
Chemistry	10
Chemical Eng.	6
Biomedical Eng.	3
Molecular Biosciences	3
Computer Science	2
Dell Medical School	1
Neuroscience	1
Civil Eng.	1
Texas Materials Inst.	1

All but one of the interviewees indicated they had external (off-campus) collaborations as well, and they named a wide variety of institutions both in the US and internationally.

In the Lab

A. Instrumentation

Chemical research is highly dependent on access to modern instruments to measure, quantify, characterize, and observe chemical processes and molecules. These instruments may be relatively small, benchtop machines that are affordable under individual grants, or larger, more expensive machines that are operated by various administrative/academic units and shared among all campus researchers (usually on a cost-recovery basis). Others are designed and fabricated in-house because no commercial versions exist. Faculty listed a wide variety of instruments in regular use, as summarized in the table.

Table 3. Types of instrumentation noted.

Instrumentation	*No. of responses*
Mass spectrometry	7
NMR	6
Other spectral techniques	6
Other various equipment	6
Microscopy (all types)	5
X-ray photoelectron spectroscopy (XPS)	4
Infrared spectroscopy	3
Lasers	3
X-ray Crystallography	3
Electron Paramagnetic Resonance spectroscopy (EPR)	2
Chromatography	2
Computers (primary)	2

Almost all instruments are digital and have dedicated computers and software to operate them, collect and manipulate data, and output results into digital file formats.

Theoretical chemists make heavy use of powerful computers as their primary instrumentation, including supercomputing facilities at TACC, to run simulations and molecular modeling programs.

Some faculty indicated that they use specialized instruments located at other institutions. One example is the Mössbauer spectroscopy facility at Texas A&M.

B. Data Gathering and Handling

There are many types of data gathered in lab settings. Spectral, kinetic, structural, crystallographic, chemometric, thermodynamic, chromatographic, and images are some of the most frequently mentioned kinds. File sizes can range from kilobytes to terabytes, depending on the source.

Instrumental output typically flows to attached computers first, where initial analysis and processing takes place under proprietary software. The paths of data after this point are very diverse: files may move to student laptops, group servers, cloud storage such as UT Box, external hard drives, even flash drives. Some groups have developed mechanisms to automatically sync multiple devices or perform backups (e.g. CrashPlan). The role of the PI in overseeing data handling also varies, from hands-on coordination to deliberately hands-off. (One chemist said, “I have absolutely no idea. It’s my students. When I see [the data] they’re in PowerPoint.”)

The management and preservation challenges created by all this data are addressed below under Data Management.

C. *Lab Notebooks and Recordkeeping*

Keeping reliable and permanent records of research activities is one of the central tenets of the scientific method. Graduate students are imbued with this requirement from the start, building on concepts learned as undergraduates. The time-honored format for these records is the laboratory notebook, in which each researcher scrupulously records notes and documentation describing the procedures and results of all their work. These notebooks serve a legal function as well, particularly in the intellectual property (IP) sphere, where the exact date and time of a particular discovery may be vital in a subsequent patenting process or dispute.

The traditional format is a bound paper notebook, written in ink, with pre-numbered pages; nothing should be erased or removed. The notebooks typically remain the official property of the PI or the institution, and in some cases may not even be removed from the lab for any reason. After they are filled they are turned in and stored centrally by the PI. In recent years a number of software developers have introduced electronic lab notebooks (ELNs) to replace the paper notebook and provide advantages such as automatic backup, versioning, direct importing of digital instrument outputs, etc. (Bogdan & Flowers 2014). ELNs are becoming the norm in industrial R&D settings where IP and centralized control of research are rigorously enforced, but they have had relatively little uptake in the more decentralized and idiosyncratic academic laboratory settings.

Thirteen of the fifteen interviewees responded to a question about lab notebooks by indicating they use only paper notebooks. A few had explored or used ELN options in the past, while the others dismissed the ELN as impractical, too expensive, too complex and time-consuming, or prone to data loss and obsolescence. (“Electronic data just seems to disappear periodically,” one said.) In 2013 the Ithaka UK chemistry report reached the same conclusion: ELNs “remain unattractive to most chemists as an option for knowledge management in the laboratory setting. This is in part because the hardware is not suited to all lab environments, but it is also because the available systems do not provide enough convenience and utility for users” (Long & Schonfeld 2014, p.12). More than five years after that study was conducted it seems ELNs have made little headway in academia. Inadequate funds, time, and IT support likely contribute to this situation (Sayre et al. 2018).

Only one professor indicated they use no notebooks at all: each lab member “just does whatever they do.”

Table 4. Recordkeeping methods.

Lab Recordkeeping Method	*No. of responses*
Paper Notebooks	13
ELN	0
Other digital	1
None mandated	1

Nevertheless, paper lab notebooks tend to be near- and middle-term solutions. Long term archival storage of filled notebooks is haphazard, and the need to retrieve old notebooks at a later date is relatively rare. They tend to be stored in boxes or bookshelves in the professor's offices, and preservation protocols are up to individual PIs. Storage elsewhere is risky: one professor reported losing many of his early notebooks during a recent cleanout of the Welch Hall attic. After a faculty member retires, dies, or leaves, the notebooks are normally discarded for lack of space. Neither the University nor the department has a retention policy specific to lab notebooks; they are not considered official records.

Information Seeking and Management

A great deal has been written over the years about how chemists find and handle information, and librarians specializing in chemical information are quite familiar with the general proclivities and habits of researchers in these fields. For that reason, we did not expect many surprises from this set of questions, and there were indeed few.

A. Primary Literature

The interview script, based as it was on a version crafted for the humanities, attempted to distinguish between "primary" and "secondary" forms of literature. These terms have a different meaning in humanities fields than they do in STEM, where they reflect a temporal position early in the research cycle. Primary literature in STEM encompasses formats such as peer-reviewed articles, dissertations, patents, conference papers, and technical reports. In retrospect this was not a useful or meaningful distinction to employ in this setting. Despite explanation, several chemists were puzzled by the distinction and may have misunderstood it. Nevertheless, some observations were clear.

Table 5. Formats of primary literature noted.

Primary Literature Format	*No. of responses*
Peer-reviewed journals	13
Patents	2
Conferences	3
Personal communications	4
Other sources	2
No response	2

If an army travels on its stomach, chemistry travels on its journals. In response to the question on what form(s) of "primary" literature they find most important for their work, all thirteen interviewees who answered this question replied that peer-reviewed journals are, and have always been, paramount. One stated simply, “We read papers.” Access to journals remains one of chemists’ key concerns, and most view this area as the most important, if not only, function of their library. Other forms of primary literature, such as patents, conference papers and dissertations, are occasionally used but regarded as less important and authoritative.

There were few explicit complaints about lack of journal access. One professor did claim that in 30 years not a single journal he had requested had been subscribed to; although this can’t be verified, it is true that the perennial lack of funds for new content can be frustrating and result in negative feelings toward the library.

B. Quality of Journal Literature

Some faculty expressed strong opinions about the perceived quality of the journal literature. Their decisions about which journals to trust, which journals they agree to review for, and where they choose to publish their own research are inextricably linked. Many commented favorably on the overall quality of the principal society-based publication programs, the American Chemical Society (now over 60 journals) and the Royal Society of Chemistry (42 titles). But they also expressed concern about what they see as a literature rife with inaccurate, low-quality, or even bogus research published in weak, non-rigorous journals. (Words like “garbage” and “junk” peppered the conversations.) This phenomenon dilutes the literature and makes it even harder to locate the good papers.

One chemist said, “It's just so much junk. The vast majority of people don't care about science anymore, they just care about getting articles in their CVs.” Another said, “I think the scientific literature just has a really basic problem in that the pressure to publish is so profound, and there's so many avenues for publishing bad stuff, that we have this gigantic foundation, much of which is just crap, useless.” On the other hand, another pointed out that “we’ve had two Nobel Prizes in our business that were based on papers originally published in … journals with an impact factor of zero, plus or minus two.”

One participant described an even more disturbing picture, where the mass of trivial literature tends to conceal the *lack* of underlying, fundamental knowledge that one would normally assume has to be there. “What we find over and over and over again is that deep questions, fundamental questions that are really important if you're trying to figure out at a molecular level how something functions, are actually very rare. And so, we'll be in a topic where I feel that there probably should be a lot of good information, and we find actually that no, that's not the case at all.” Tracking citations backwards, the professor claimed, leads to the realization that “the literature is just a mess, and that we're really going to have to do [the research] from first principles.”

C. Discovery

Most faculty described ad-hoc approaches to literature searching within their groups. Everyone spends time doing it, and papers of interest are shared around informal networks or presented in group meetings. Some search on a regular basis, as a means of keeping up, while others search only when the need arises, especially when exploring an unfamiliar topic, preparing grant applications, and writing papers. Group members are typically free to choose their own tools and methods; one faculty member said simply that “everybody does their own thing.”

The selection of search tools seems to be driven primarily by comfort and familiarity, and once habits are formed they are not easily changed. Simplicity emerged as an overriding preference. When interviewees indicated they had changed preferred tools in the past, it was usually a move to a less complex tool (e.g.

from SciFinder to Google Scholar). The more complex, feature-laden tools such as SciFinder are more often criticized as unwieldy or unfriendly. When asked about SciFinder (the platform for the Chemical Abstracts and Registry databases), one professor remarked, "We used to use it all the time, but…I've almost given up on it. And I don't think any of my students use it now." Google Scholar and Web of Science are favored for most text-based tasks, including author searching (see Gordon et al. 2018). When the need arises to do more specialized "chemistry-based" searching (using drawn chemical structures and reactions, precise chemical nomenclature, or - less frequently - numeric data values), SciFinder and Reaxys are the only major literature discovery tools that offer these powerful features. Some faculty indicated that it's mostly their graduate students who use these types of databases, which have a steeper learning curve and require more regular practice to use proficiently. One professor who doesn't use SciFinder himself acknowledged this: "I've sat next to people when they were using SciFinder, people whose research is very structure-driven, and it's amazing how quickly they can find things and do things." A database's scope is also a point of distinction. The popularity of Web of Science may stem from its focus on a curated list of "better" journals, as opposed to the more comprehensive scope of SciFinder or the black-box coverage of Google Scholar.

Table 6. Databases and search engines noted by priority.

Discovery Tools	*Primary*	*Secondary*
Web of Science	5	2
Google	3	2
SciFinder	2	5
Google Scholar	2	2
None specified	2	0
PubMed/MEDLINE	1	2
Reaxys	0	1
Inspec	0	1
USPTO	0	1

D. Secondary Literature

STEM librarians use the term "secondary literature" to designate monographs, edited books, and review articles, all of which serve to summarize and synthesize current knowledge of a topic rather than reporting new empirical results. The chemists interviewed made little distinction between review articles and primary research articles since they are published side by side in journals and are discovered in the same ways. Likewise, the distinction between research-level monographs and pedagogical textbooks is largely meaningless to them. Several interviewees described occasional use of books, but primarily for background information and teaching, not for their research per se. Interest in books seems to be dropping, however: one remarked that "we have really stepped away from using books." Another said, "It's weird that I'd look into a book for a research topic," and a third stated, "Usually we don't even bother looking for them."

While some indicated they use library copies and prefer print books over electronic, another mentioned the difficulties encountered in discovering books, the contents of which are generally not represented in the main indexing services and which are more difficult to obtain even if one knows about them. Google and Amazon might serve as discovery mechanisms, but they provide no access to the full text in most cases. Only one interviewee mentioned the Library Catalog as a discovery tool in this context, but three stated that they liked to browse in the library stacks, and lamented the recent closure of the Chemistry Library, which ended their ability to do that. Finally, one chemist marveled at the ability of his graduate students to obtain perfect PDF copies of books "from China" – presumably pirated – almost instantaneously. "I suppose I should tell them that this is unethical, but it's so convenient anyway."

If faculty tend to have low opinions about the relative value of monographic materials, this is reflected in their use of the library. Most books checked out from the library go to graduate students and postdocs, not faculty. (In 2016 faculty patrons accounted for only 7.6% of books checked out from the Chemistry Library's circulating collection.)

E. Organization and Management

Organizing and managing the mass of PDF articles collected over months and years of searching and sharing is a serious challenge, and many of those interviewed expressed frustration at the lack of adequate tools to make this task easier. Some have developed personalized filing systems on their own computers, using software such as Endnote or Zotero, while others have tried external tools such as Papers, Readcube, and Mendeley. Most, however, simply file away PDFs into folders. This makes it difficult to find specific papers later – searching for and downloading a fresh copy is sometimes easier than locating an earlier copy that may or may not have been saved. Surprisingly, four interviewees said they prefer to print out and file paper articles. Asked about neat stacks of articles arrayed on his desk, one professor confessed, "The ones you see out are the ones I haven't read. Which is very sad." It's also clear that within lab groups everyone is on their own figuring out these questions; no faculty member described any form of established procedures or central repository for gathering, storing, and sharing articles among group members.

F. Keeping up with the Literature

The sheer volume of new literature makes it nearly impossible to keep up with trends and developments in one's own research area, let alone in a wider sphere of interest. As with point-of-need discovery, faculty members' coping strategies are ad-hoc and tend to be hit or miss. In the print era, making time to visit the library and browse new issues of journals in the reading room was part of the normal routine for some faculty. One recalled: "When we had a library that had active subscriptions to journals, I would go down there every week and go through about 50 journals. The library over time quit subscribing to hardcopies…, [s]o I can't do that anymore." Journals can of course be browsed on the Web, but it turns out the computer is slower than the eye, as the same chemist pointed out. "Frankly, browsing even graphical abstracts is painful online. It's so much easier to take a hard copy of a journal and browse than it is to browse on a computer. So, honestly, I think we've taken a step back in digitizing everything."

Strategies may be group-oriented: several interviewees indicated they rely on their students to bring them papers of interest. A few said they use "keep me posted" alert features in various databases such as Google Scholar or Web of Science, but this doesn't seem to be a widely used method. Others sign up for email table-of-contents services from their favorite journals, but some have given up on that tactic as too voluminous and unfocused, as well as self-limiting.

Most interview subjects pointed to attending conferences as an important way to learn about new research even before it's published. These could be large conferences such as twice-yearly ACS national meetings where thousands of formal presentations and posters are given. A chemist described a divide-and-conquer approach: "I really rely on my graduate students, who act like a little army. If we go to a conference, they can all cover a lot more ground than me, and they have their own specific niches that they can map out. So I rely on them to go out, find out what's hot, what's new, and then report back after the conference." Others prefer smaller meetings, such as Gordon Conferences, as more focused and manageable than the giant conventions. Speaking at invited seminars at other institutions is also a popular tradition that allows conversations and networking in more intimate settings.

Some acknowledge that keeping up is a losing battle. One professor rationalized his relative indifference to current awareness methods: "My experience is the more you know, the less likely you are to be creative, because you look at something and you just talk yourself out of good ideas." Knowing more can apparently come from knowing less.

G. External Data Needs

When asked about the need for obtaining research data that they do not generate themselves, most of the interviewees responded that they rarely or never need such data. Some make use of established specialized repositories such as the Cambridge Structural Database (crystallographic data of organic substances) and the Protein Data Bank. In chemistry as in most hard sciences, processed datasets that underlie peer-reviewed papers, but which are too large to be included within the paper itself, are usually provided as Supporting Information alongside associated journal articles. Scientists regard these as important sources for any thorough reading of reported research, even if the data have been converted to flat-file text or PDFs (Martinsen 2017). Beyond these examples, the consensus among faculty interviewed was that the need to see others' raw data is relatively uncommon.

One interviewee mentioned dissertations, an often-ignored form of primary literature in chemistry, as a source of data. "I definitely find myself searching for dissertations every now and then. I work in a field in which people who are publishing tend to generate very large data sets, which often don't make it into a paper. There's some observation that makes it into a paper, but you need the data, and so we'll … hunt down a dissertation … because I was told that the PhD dissertation had a lot of data tables in it that we wanted to use."

Publication and Dissemination Practices

A. Venue Selection

The rule of thumb for publishing one's own research is that PIs aim for the journal with the highest perceived quality that they can reasonably hope will accept their paper. The intended audience is also an important factor, depending on the topic. There is a strong bias in favor of the American Chemical Society journal program, which almost every interviewee cited as the top destination in chemistry and related fields. The Royal Society of Chemistry also received a number of favorable mentions, as did *Proceedings of the National Academy of Sciences* (PNAS). Conversely, some indicated a bias against commercial, for-profit journals, with a few notable exceptions such as Wiley's *Angewandte Chemie Int. Ed.* and upper-tier titles in the Nature portfolio. One professor expressed it clearly: "It's a combination of where we're going to reach the right audience, and how do we reach the largest audience? Which aren't always the same. How do we support more nonprofit, academically oriented publishing ventures versus lining the pockets of big companies? So certainly, all things equal, we absolutely prefer to publish in an academically oriented place rather than one of the big dogs."

To corroborate these claims, we used Web of Science to analyze the 579 articles co-authored at UT-Austin by the 15 interviewees during the time period of 2013 to mid-2018. Of 66 journals with two or more papers, 34 (52%) were indeed from ACS or RSC, together accounting for 360 (70%) of the 517 papers published in these 66 journals (Table 7; see also Appendix 2, Table 1). Wiley was the top commercial publisher in this group, with 57 papers (11%). Their highly regarded *Angewandte Chemie Int. Ed.* was the highest-ranked for-profit journal in the list, with 17 papers.

Table 7. Journal titles with 2 or more articles co-authored by interviewees, 2013-18, by publisher.

PUBLISHER	*Number of Journal titles*	*Number of Articles*
ACS	21	257
RSC	13	103
WILEY	10	57
ELSEVIER	9	30
AIP	3	30
*NATURE**	3	13
OTHER	7	27
TOTAL	66	517

Source: Web of Science
** Includes Nature-branded titles only*

A few faculty members specifically objected to an increasingly common model of manuscript processing called “journal cascades,” whereby a publisher seeks to retain in their control manuscripts that don’t make the cut at the intended journal. It is passed downwards through a portfolio of “lesser” titles until it is accepted. The business case for this model is simple: handling, vetting and reviewing manuscripts is costly, and if a paper is rejected it might otherwise start all over with a rival journal, and eventually it’s almost certain to be accepted somewhere. The solution is to capture these papers in-house so that the upfront investment is not wasted, and the authors are saved the time of resubmitting elsewhere. But when journals at the lower levels of the cascade charge author fees, sometimes without the authors’ prior consent, the process becomes more fraught. One chemist was particularly annoyed by this practice, and coined the term “publication pachinko” to describe it. Another PI told of withdrawing a paper that had been passed down to *Nature Communications* without his approval, rather than pay an author fee of over $5,000 to publish it there. “I'll send it somewhere else, thank you, and save my money,” he told them.

Serving as a peer reviewer for journals is also a part of the PI’s duties. Some faculty deliberately limit the number of journals they publish in in order to reduce the demand for reviewing, which can be very onerous and time-consuming, especially for senior faculty. It’s easier to say no to journals that you don’t publish in, as one chemist related: “I'll just say, ‘Well, I'm sorry, we don't submit to you, so I'm not refereeing for you.’ That's the only way to cut down, otherwise you'd spend your whole life refereeing. I'll referee and never turn down an invitation or a request to referee something in a journal that we publish in. With all the others it's just, ‘Sorry, not doing it.’"

B. Open Access

Chemists have long had a reputation for indifference when it comes to the "Open" movement. Chemistry as a discipline lags all other STEM fields in making the research literature available without paywalls (Piwowar et al. 2018). The faculty members interviewed for this study displayed a fairly nuanced view of the open question, but for the most part it is simply not a priority for them at this time. Almost all of them denied having ever paid an author processing charge (APC) to make an article open in a Gold or hybrid OA journal, and said they would never normally choose this option. While none objected to the notion of a literature that's free for all to read, some feel that it's not a major problem and certainly not something they think they should pay for directly out of their grants. A selection of quotations illustrates this attitude:

"It's just not an issue at all. My audience are people at places like UT. I'm not worried about people who don't have access to those journals."

"I think [paying APCs] is a waste of money…. I am opposed to paying a fee if there's something that I can upload to another site myself or if it just has to wait a year, then so be it."

"The presence or absence of a journal's open access policy basically has no bearing on whether or not we would choose to submit…to that journal."

"I'm a firm believer in Open Access, but I guess not to the extent of paying a thousand bucks per article."

"Personally I don't think authors should have to pay to publish their stuff. Publishers make plenty of money."

"I think it's a neat idea that you can just let anyone have access to your results."

They expressed a similar disregard for Green OA deposit of post-prints, with the notable exception of compliance with NIH mandates to deposit manuscripts in PubMed Central after 12 months. We did not ask them specifically about their awareness or use of UT Austin's local repository, Texas ScholarWorks, but none of them mentioned it or implied they had heard of it. That said, some chemists take the Green OA concept into their own hands: a few confessed that they post final publisher PDFs of their articles to their personal lab web sites, regardless of copyright legality or concern about the consequences. "I guess we probably shouldn't be doing that," one admitted, while another was more defiant: "Sometimes the legality of that is a little bit in that gray area, but to be honest, I don't care. I think that having open journals is the most important thing. The same thing with book chapters."

Unlike physics, chemistry has no tradition of posting or sharing preprints in advance of peer-reviewed publication. Only one faculty member mentioned having submitted a preprint to arXiv, the dominant repository in physics and mathematics. There have been some recent attempts to jump-start a preprint culture in chemistry, but these are not yet widely known.[1] Chemists seem to prefer to wait for peer review to play out, providing the opportunity to improve an article and correct errors before it's publicly available. One chemist explained his concern about the persistence of multiple versions of an article: "You can never get rid of it. And so now, you've got two articles out there, version one and version two, and there's an error in version one that a peer reviewer called your attention to and they get fixed in the published article. But if somebody does a web search, who knows which one's going to pop up."

While the faculty interviewed tended to express indifference or hostility to various Gold-OA mechanisms, a bibliometric analysis of their papers provided an interesting counterpoint. The Web of Science search retrieved 579 articles where the interviewed PIs were listed as co-authors, published between 2013 and mid-2018. Of these, 193 (33%) linked to open versions of some kind. (See Appendix 2, Table 2.) The majority of these (111) were classified as "gold or bronze" meaning that a fee had been paid by someone, while 82 were classified as "green" and located in a repository, most commonly PubMed Central. Since articles are a group effort, sometimes crossing departmental and institutional boundaries, some co-authors may not always be aware of an article's open status, so it's understandable that they may believe it happens less than it does. Still, at a time when some disciplines are approaching a 50% open rate, chemistry (at around 20% according to Piwowar et al.) is clearly behind the curve.

In the literature-centric culture of chemistry, ingrained biases and the existing reward system militate against significant change on either the Gold or Green OA fronts. Chemists are indoctrinated through the graduate school/postdoc apprenticeship system to believe in the efficacy of the existing pecking order of established journals, virtually all of which adhere to the pay-to-read model. As they reach the tenure track, the career reward system in place within higher education, which values not only quantity of publications but also the perceived quality of the journals, reinforces the status quo by disincentivizing the exploration of alternative publication outlets. This lesson is then passed down to the next generation of researchers in a cycle that is very difficult to break.

So far, few universities even mention OA in promotion/tenure processes, and those that do often do so in a negative way, out of fear that "open" may equate with "predatory" or low quality (Alperin et al. 2018). The plague of spam from fraudulent and disreputable "publishers" reinforces this view.

A chemist explained it this way: "There's a tremendous pressure, particularly [on] the younger members of the community, to publish in these so-called 'high impact' journals. In fact, [an administrator] last year did not help this, or exacerbated it, or played into this, by sending around a memo, 'Please list your high-impact papers for 2016-2017.'" What the administrator meant by "high impact" in this context was left undefined.[2]

Academic scientists see themselves as independent entities and are skeptical when others attempt to dictate how and where to publish. Open mandates, which represent added time, work, and costs, are only grudgingly accepted when they are required (and enforced) by funding agencies, the most important of which is NIH (Van Noorden 2014). Most other mandates, especially those that are not driven by the research community itself, tend to lack enforcement mechanisms and are more easily ignored. For example, chemistry papers funded by NSF, which has a weaker mandate, showed only a 24% open rate, compared with 81% for NIH papers (Larivière & Sugimoto 2018).

---

[1] ACS' ChemRxiv and Elsevier's ChemRN both launched in mid-2017.

[2] In response to a later follow-up inquiry, the professor indicated that this request "came with no instructions" and he personally decided to list papers published in journals with an impact factor ≥10.

C. Dissemination and Outreach

The faculty interviewed mentioned three primary approaches to publicizing their research beyond merely publishing articles. But they aren't equal in terms of uptake.

1. Presenting at **conferences, symposia, and seminars** was mentioned most often – many PIs do this routinely as a way to talk about their work and connect with others who share their interests. This is often how collaborations start.
2. **Lab group web sites** are a standard venue for listing (and sometimes posting) publications, occasionally with graphical abstracts. These pages also serve as a platform for summarizing research areas and highlighting the achievements and activities of individual group members. Faculty find these sites valuable both for promoting their own work and tracking the work of their colleagues and competitors.
3. Only one interviewee noted the use of **social media**, specifically Twitter and LinkedIn, as a mechanism for publicizing and discussing recent work. (A second interviewee who has Twitter and LinkedIn pages was not asked this question.) Several others specifically said they don't use social media, although some allowed that their students might. "We publish and then, for me anyway, that's the end of it," one remarked.

Other "reputation management" tools are trending today, but none of the interviewees brought up ORCiD, ResearcherID, ResearchGate, Google Scholar Citations, or other such platforms in this context, and we didn't ask about this topic specifically.

To follow up on this question, we performed Google searches on each interviewee (firstname lastname Texas chemistry) and examined the first three pages of results, noting which outreach platforms came up (Table 8). As expected, all 15 had up to date research group web sites that appeared near the top of the search results. Google Scholar Citations was a distant second with eight faculty having set up publications pages. LinkedIn and ResearchGate were also in the mix (stub pages weren't counted), but no hits on similar services such as Mendeley or Academia.edu turned up. Twitter and Facebook were only counted if the faculty member used it for professional/research dissemination.

Table 8. Faculty Web and Social Media Presence via Google search.

Platform	*No. of faculty*
Lab/Research Web Site	15
Google Scholar Citations	8
LinkedIn	6
ResearchGate	3
Twitter	2
Facebook	1

Data Management

“We may not do so well.”

This quote from one chemist encapsulates the dilemma facing researchers today as they consider ways to manage, preserve and share the vast amount of data they generate on a daily basis. It was clear from the majority of interviews that data management strategies within lab groups are ad-hoc, underdeveloped, or nonexistent. This is not always for lack of awareness or effort, but PIs are greatly handicapped by the lack of adequate, appropriate tools for the job. They must often resort to cobbling together makeshift solutions with off-the-shelf hardware or cloud storage systems such as Box which aren’t ideal for the purpose. Every component of data storage and management is beset with challenges: server/hard drive capacity, network speed, security flaws, software maintenance, redundancy, automated backup, IT support, data loss, and more. The more numerous and voluminous the data files are, the greater the challenges in moving, sharing, and preserving them. The diversity of file types and data formats, especially related to output from instruments of various purposes and ages, creates further hassles. The whole situation is, as one faculty member put it, “a major headache.”

Data can take the form of numeric files, text files, static or moving images, structural data, and more. Specialized software or code, sometimes custom developed in-house, is usually needed to process and visualize instrumental output. Raw output without accompanying code may be useless. The quantity of data also varies considerably. Some instrument output is at the megabyte level, but other output, especially high-speed moving images, micrographs, and the like, rapidly accumulate at the terabyte level, far too large to be effectively stored in cloud systems like Box or in open data repositories.

While some professors attempt to maintain a centralized and coordinated storage system across the whole group, hands-off ‘every-person-for-themselves’ approaches seem to dominate. One professor stated that his “students store all their data on their personal laptops. … We have talked numerous times about having some kind of cloud source where we put everything, and you probably remind me I should talk about it at group meeting.” Another was equally vague: “I know my students have some archival data. I personally don’t manage any of that. … Who knows where the data is. It’s probably on their computers.” Still another puts the onus on his students: “I expect my students to maintain their data on their own computer and then on some kind of backup.”

Faculty also question the need to devote time and resources to the long-term preservation of raw data that may never be needed again. “I might know we have to keep it. We’re doing our best. But years go by and nobody requests that data,” one said. In some cases, preserving and later finding old data is less efficient than starting from scratch. “There’s nothing that we do that can’t be recreated if we need to,” said another. A third stated: “We don’t really have a use for long-term storage of data. [T]he raw data itself, I don’t really think that’s worth preserving…. Looking at a dataset that’s old, in my opinion, has less value than collecting a fresh dataset. [W]e rarely archive data for more than a year.”

They also drew a distinction between raw and processed data. Underlying data that are published within a journal article, or made available as supporting/supplementary information in conjunction with an article, or even just referred to in a publication, are clearly important to them; but raw, unprocessed data such as instrument output are seen as disposable.

These responses belied the recent regulatory trend that mandates "data management plans" (DMPs) as part of grant applications. None of the faculty interviewed mentioned a DMP specifically. Two professors who alluded to the data preservation rules, particularly from NSF, indicated that the definition of data storage and sharing, as they see it, is fairly broad, and they expressed confidence that whatever they're doing now meets the standard.

Because the interview was deliberately constructed to focus on the researchers' processes rather than the library per se, we did not ask the faculty if they envisioned any role for the library in the data management equation. Absent that prompting, none of the faculty indicated in their remarks that they saw such a role.

Data Sharing

The chemists interviewed tended to express hesitation when asked about sharing their data with others. One pointed out a fear of being scooped or giving away potentially valuable information. Another interviewee requires industrial collaborators to sign a nondisclosure agreement prior to receiving any data. A third was less guarded: "I'll give it to anyone who wants it, and I definitely have." But for the most part, the faculty interviewed simply don't think that their internal data – particularly the unprocessed, unpublished versions – have much value beyond their own labs, and therefore they see little point in expending the effort to preserve them in open repositories for anyone to use. Instead they respond to direct requests from other scientists, and evaluate them on a case by case basis. One professor said he posts datasets on his personal lab web site, but lamented that, apart from Google, there was no way to find them and they lacked direct linkage to published papers. Faculty were not asked about their awareness of the recently launched Texas Data Repository (TDR). One interviewee mentioned it without prompting, but said she had not used it.

There is greater understanding and acceptance of sharing in a few specific areas where deposit in established (but not necessarily open) disciplinary repositories is already the norm: crystallographic data files (e.g. Cambridge Structural Database), protein structure data (e.g. Protein Data Bank), proteomic data (e.g. JPOST), software code (Github, Figshare), etc.

A Magic Wand?

The interviews concluded with a hypothetical question that invited the subjects to speculate on what they might wish to see if they could "wave a magic wand" and improve some aspect of their work lives in terms of information and data – apart from gaining more money or time. The usefulness of responses varied, since people often don't have such a wish-list at their mental fingertips. Some of the core responses, in no particular order:

- Better literature search engines, possibly using AI technology to analyze and filter results and navigate the information stream. (There were several variations on this particular theme.)
- Fewer journals to read and track.
- Better way to filter out "garbage" science.
- Better method of getting literature alerts for new papers or citations.
- Better way to manage and archive downloaded papers.
- Better IT infrastructure on campus that more directly supports researchers, e.g. dedicated servers for data backup, faster networking speeds, better IT support.

- Reduced obligations to review papers for journals.
- Return to hardcopy submission process for journals.
- End the “journal cascade” model within publisher portfolios.

If there was an overarching takeaway to these responses and to a follow-up question on the general challenges facing the discipline, it was that the younger, newer faculty seem more optimistic about the future, while the more senior faculty tended to express more feelings of burnout; indifference to new technologies; exasperation with administrative and bureaucratic obstacles; helplessness in the face of a relentless blast of new literature and editors’ demands for ever-faster turnaround times; and more resistance to change.

Overall this is not surprising. Scientists with long careers have seen a steady erosion of resources, infrastructure, and administrative support, all while the burdens of compliance, higher costs, more complex technology, and expectations of productivity have only grown more daunting. One chemist who is nearing retirement said simply, “I'm [age redacted] years old and I don't have an iPhone; I have a flip phone and I'm just not willing to invest an enormous amount of time learning how things work. I would rather focus on science.”

REFERENCES

Alperin, J.P., Muñoz Nieves, C., Schimanski, L., Fischman, G.E., Niles, M.T. & McKiernan, E.C. (2018). "How significant are the public dimensions of faculty work in review, promotion, and tenure documents?" Humanities Commons [preprint]. DOI: http://dx.doi.org/10.17613/M6W950N35

Bogdan, K. and Flowers, T. 2014. "Electronic lab notebooks: supporting laboratory data in the digital era." *Issues in Science and Technology Librarianship*, Spring. http://www.istl.org/14-spring/app.html

Flaxbart, D. 2003. "Conversations with chemists: information seeking behavior of chemistry faculty in the electronic age." *Science & Technology Libraries*, 21(3/4), 5-26. http://hdl.handle.net/2152/63911

Gordon, I., Meindl, P., White, M., and Szigeti, K. 2018. "Information seeking behaviors, attitudes, and choices of academic chemists." *Science & Technology Libraries*, 37(2), 130-151. DOI: https://doi.org/10.1080/0194262X.2018.1445063

Larivière, V. and Sugimoto, C.R. 2018. "Do authors comply when funders enforce open access to research?" *Nature*, 562, 483-486. DOI: https://doi.org/10.1038/d41586-018-07101-w

Long, M.P. and Schonfeld, R.C. 2013. *Supporting the changing research practices of chemists: research support services: Chemistry project.* (Ithaka S+R.) http://www.sr.ithaka.org/publications/supporting-the-changing-research-practices-of-chemists/

Martinsen, D. 2017. "Primary research data and scholarly communication." *Chemistry International*, 39(3), 35-38. DOI : https://doi.org/10.1515/ci-2017-0309

Piwowar, H. et al. 2018. "The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles." *PeerJ*. DOI: https://doi.org/10.7717/peerj.4375

Sayre, F., Bakker, C. and Kelly, J. 2018. "Support for electronic lab notebooks at top American research universities." *Journal of eScience Librarianship*, 7(2), article 8. DOI: https://doi.org/10.7191/jeslib.2018.1140

Van Noorden, R. 2014. "Funders punish open-access dodgers." *Nature*, 508(7495), 161, DOI:10.1038/508161a

ACKNOWLEDGEMENTS

We appreciate the support and assistance of other library staff in conducting this study: Bonnie Brown-Real, Krystal Wyatt-Baxter, and the staff in the Captioning Service unit.

APPENDIX 1

Semi-Structured Interview Guide

Research focus and methods

1. Describe your current research focus/projects.
    a. Did you pursue funding for these projects? Tell me about that process.
    b. Are you collaborating with others at UT Austin? From another institution?
    c. What types of equipment or instrumentation do you rely on to conduct your research?
2. What kind of data output, if any, does the research generate? [If so, probe for how they manage and store this data for their ongoing use or sharing]
3. How do you share, organize, save, research notes within your group?
    a. Does your group use an electronic lab notebook system (ELN)? [If yes, which one, and how's it working?]

Information Access and Discovery

1. What kinds of primary information do you rely on to do your research? [Examples: peer-reviewed journal articles; patents; personal communications; etc]
    a. How do you locate this information?
    b. What are the greatest challenges you experience working with this kind of information?
    c. How do you manage and store this information for your ongoing use?
2. Do you typically need to locate **data** (other than data you generate yourself) to support the planning and carrying out of your projects? [If yes, probe for what kind of data, formats, and where it's obtained]
3. What kinds of **secondary** information do you rely on to do your research? E.g. monographs, reviews, edited books, databases, etc.
    a. How do you locate this information?
    b. What are the greatest challenges you experience working with this kind of information?
    c. How do you manage and store this information for your ongoing use?
4. Think back to a past or ongoing research project where you faced challenges in the process of finding and accessing information.
    a. Describe these challenges.
    b. What could have been done to mitigate these challenges? [Or what did you do to mitigate these challenges?]
5. How do you keep up with trends in your field more broadly?

Dissemination Practices

6. Where do you typically publish your research? [Probe for kinds of publications and what disciplinary audiences they typically seek to engage with]
    a. Do you disseminate your research beyond scholarly publications? [If so, probe for where they publish and why they publish in these venues]
    b. How do your publishing practices relate to those typical to your discipline?
7. Have you ever made your research data, materials or publications available through open access? (e.g. paid a fee for a paywalled journal, deposit in an online repository)
    a. If so, where, and what have been your motivations for pursuing open dissemination channels? (i.e. funder-required, for sharing, investment in open access principles)
    b. If no, why not? [Would you share this data with other researchers if they asked?]

8. How do you store and manage data produced in your lab for the long term? (ie, after publication, after students graduate, etc)

State of the Field and Wrapping Up

9. If you had a magic wand that could help you with your research and publication process [except for more money or time] – what would you ask it to do?
10. What future challenges and opportunities do you see for the broader field of Chemistry and STEM in general?
11. Is there anything else about your experiences that you think we should know that was not covered in the previous questions?

APPENDIX 2: Bibliometric Tables

Table 1. Journals with 2 or more articles by interviewed authors, 2013-2018. (Source: Web of Science)

Source Titles	*Publisher*	*# records*
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY	ACS	65
ANALYTICAL CHEMISTRY	ACS	55
CHEMICAL COMMUNICATIONS	RSC	31
JOURNAL OF CHEMICAL PHYSICS	AIP	26
*CHEMICAL SCIENCE**	RSC	23
ANGEWANDTE CHEMIE INTERNATIONAL EDITION	WILEY	17
INORGANIC CHEMISTRY	ACS	16
LANGMUIR	ACS	16
JOURNAL OF PHYSICAL CHEMISTRY C	ACS	15
CHEMISTRY A EUROPEAN JOURNAL	WILEY	13
JOURNAL OF PHYSICAL CHEMISTRY B	ACS	12
JOURNAL OF PHYSICAL CHEMISTRY LETTERS	ACS	12
DALTON TRANSACTIONS	RSC	11
JOURNAL OF THE AMERICAN SOCIETY FOR MASS SPECTROMETRY	WILEY	11
ORGANIC LETTERS	ACS	11
TETRAHEDRON	ELSEVIER	11
PHYSICAL CHEMISTRY CHEMICAL PHYSICS	RSC	10
ACS NANO	ACS	9
ACS CATALYSIS	ACS	8
NATURE CHEMISTRY	NATURE***	8
PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA	NAS	8
CHEMISTRY OF MATERIALS	ACS	7
*SUPRAMOLECULAR CHEMISTRY***	T&F	7
JOURNAL OF PROTEOME RESEARCH	ACS	6
LAB ON A CHIP	RSC	6
JOURNAL OF MATERIALS CHEMISTRY A	RSC	5
ACS APPLIED MATERIALS INTERFACES	ACS	4
ACS MEDICINAL CHEMISTRY LETTERS	ACS	4
ANALYST	RSC	4
BIOORGANIC MEDICINAL CHEMISTRY	ELSEVIER	4
JOURNAL OF ORGANIC CHEMISTRY	ACS	4
*JOURNAL OF PORPHYRINS AND PHTHALOCYANINES***	WORLD SCI	4
ACS CHEMICAL BIOLOGY	ACS	3

*CHEMELECTROCHEM***	WILEY	3
INTERNATIONAL JOURNAL OF MASS SPECTROMETRY	ELSEVIER	3
MOLECULAR MICROBIOLOGY	WILEY	3
RSC ADVANCES	RSC	3
*SCIENTIFIC REPORTS**	NATURE	3
ACS CHEMICAL NEUROSCIENCE	ACS	2
ACS SENSORS	ACS	2
ADVANCED ENERGY MATERIALS	WILEY	2
ADVANCED FUNCTIONAL MATERIALS	WILEY	2
APPLIED SURFACE SCIENCE	ELSEVIER	2
BIOCHEMISTRY	ACS	2
BIOORGANIC MEDICINAL CHEMISTRY LETTERS	ELSEVIER	2
CANCER RESEARCH	AACR	2
CATALYSIS SCIENCE TECHNOLOGY	RSC	2
CHEM	ELSEVIER	2
EUROPEAN JOURNAL OF ORGANIC CHEMISTRY	WILEY	2
GENES DEVELOPMENT	HIGHWIRE	2
JOURNAL OF APPLIED PHYSICS	AIP	2
JOURNAL OF CHEMICAL EDUCATION	ACS	2
JOURNAL OF MATERIALS CHEMISTRY B	RSC	2
JOURNAL OF MEDICINAL CHEMISTRY	ACS	2
JOURNAL OF THE ELECTROCHEMICAL SOCIETY	ECS	2
*MBIO**	ASM	2
NANOSCALE	RSC	2
NATURE COMMUNICATIONS	NATURE	2
NEW JOURNAL OF CHEMISTRY	RSC	2
ORGANIC BIOMOLECULAR CHEMISTRY	RSC	2
PHYSICAL REVIEW B	AIP	2
PROTEIN SCIENCE	WILEY	2
PROTEOMICS	WILEY	2
SURFACE SCIENCE	ELSEVIER	2
SYNTHETIC METALS	ELSEVIER	2
TETRAHEDRON LETTERS	ELSEVIER	2

* Open access title
** Unsubscribed title
*** Nature-branded title

Table 2. Open Access articles by publication year. (Source: Web of Science)

Publication Year	*# Records*	*Percent*
2018 (Jan-Jun)	7	3.6%
2017	30	15.5%
2016	46	23.8%
2015	43	22.3%
2014	35	18.1%
2013	32	16.6%
Total	193	100%