

8-30-2010

# Trends in Chemical Information Literacy and Collection Development, 2000–2009

Jeremy R. Garritano

*Purdue University*, [jgarrita@purdue.edu](mailto:jgarrita@purdue.edu)

Follow this and additional works at: [http://docs.lib.purdue.edu/lib\\_research](http://docs.lib.purdue.edu/lib_research)



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

---

Garritano, Jeremy R., "Trends in Chemical Information Literacy and Collection Development, 2000–2009" (2010). *Libraries Research Publications*. Paper 152.

[http://docs.lib.purdue.edu/lib\\_research/152](http://docs.lib.purdue.edu/lib_research/152)

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

## **Trends in Chemical Information Literacy and Collection Development, 2000-2009**

**Jeremy R. Garritano**

Address correspondence to Jeremy R. Garritano, Chemical Information Specialist and Associate Professor of Library Science, Purdue University Libraries – CHEM, 504 West State St., West Lafayette, IN 47907, USA. E-mail: [jgarrita@purdue.edu](mailto:jgarrita@purdue.edu)

### **Abstract**

This review covers major trends in the field of chemical information over the last ten years and how they influenced information literacy and collection development practices of chemical information specialists. Particular attention is given to changes in information literacy and discovery practices as they relate to developments in chemical information resources, the integration of various resources online, and the creating of information literacy standards in the 2000s. Also, changes in licensing and purchasing for online resources as they relate to pricing models, ownership, platform, and distribution are discussed as major influences to new collection development practices.

## **Introduction**

The author received his Master's in Library Science in May 2000. So a review of major trends affecting chemical information in the last decade seems fitting as the years 2000 to 2009 basically bound the entirety of the author's professional career. It is remarkable that even in this short time, major changes in the field of library and information science have greatly affected the approaches the author has taken toward his work as an information professional. In the last ten years the author has worked at a mid-level research university as a general science librarian, then as a visiting science librarian at a small liberal arts college, and for the last six years has been a chemical information specialist at a major research university. The author's major responsibilities within each of these positions have focused on areas related to information literacy and collection development. Consistently participating in these two areas of library and information science provide the common threads of analysis for the following work. Looking at how these two areas have changed over the last decade and how the author's own perspectives and behaviors have changed will demonstrate the critical issues facing most chemical information specialists in academic libraries during the last ten years.

While concepts that will be discussed such as information literacy, substructure searching, and resource licensing have all existed prior to 2000 it is mostly due to the rapid advancement and availability of the Internet that has caused these concepts to be radically changed. This in turn caused librarians to reevaluate the decisions they made related to instruction, information discovery, and collection development. While not an exhaustive list, the issues discussed here were major influences in the author's career as it developed over the decade.

## **Changes in Information Literacy and Discovery Practices**

Over the last ten years, once familiar chemical information resources were increasingly being adapted to the online environment. Whether as stand-alone CD-ROMs, client software, or online resources to be accessed over the Internet, these new resources changed the information discovery process, and therefore the methods by which information professionals taught these resources to users. These changes can be discussed by examining how improvements to existing resources (in terms of both features and platform), integration between resources, and the creation of entirely new information resources influenced information literacy and discovery. Related to these changes was the evolution of information literacy standards created by various organizations. These standards sought to provide guidance in a dynamic information environment and to provide the information literacy skills that future chemists needed, despite changes in tools and platforms.

### **Improvements to Existing Resources: Features**

Perhaps one of the most commonly discussed and most important chemical information resources, *Chemical Abstracts*, was first provided as client software to academic institutions in 1998 (Somerville 1998). By 2000, the core of the client software, named SciFinder Scholar (SFS), was fully established and included the ability to perform searches by structure, substructure, reaction, author, topic, and a few other choices, even if it was limited to searching the literature from 1967 to the present (Bolek 2000). Each year usually brought one or two enhancements to the software. While this required extra work on a library's part to distribute and reinstall each new version, the enhancements often made up for the inconvenience. Initially the primary searches that could be conducted via SFS were similar enough to the entry points used to search *Chemical Abstracts* and it became simply a matter of helping users adjust from one format to another. And to a certain extent, SFS itself helped with this adjustment by freeing

the user from the more rigid controlled vocabulary required to search *Chemical Abstracts*. So, while a user could perform an author search in either *Chemical Abstracts* or SFS, SFS could help users find misspellings and alternative spellings of author names. SFS could further extend subject and structure searches by retrieving synonymous terms or finding substructures within molecules. Thus, a trend toward more natural language or ambiguous searching was provided to library users.

However, in February 2000, the next iteration of SFS that was released would demonstrate the true power of an indexing and abstracting service in an online environment. The February 2000 enhancements included a new “analyze references” feature, among others. This new feature allowed users to analyze their results in a way they could not by simply retrieving a list of disjointed references from multiple print volumes—they could now have SFS create histograms giving them the most frequent authors within a hit set or list the organizations most involved in a particular area of research. This feature allowed users to identify overall trends for a particular topic or within a particular author’s body of work and analyze the data itself compared to simply dealing with discrete references. While this could possibly be replicated by hand using *Chemical Abstracts*, SFS saved a tremendous amount of time and offered users a competitive intelligence tool that had not readily been available to them before. Teaching users how to use these features in new ways or having them consider the different purposes for these features became crucial for librarians to understand and disseminate. While the interface did not change drastically, each new release of SFS had new features that merely appeared as additional buttons or sub-menus within the program, and these features were not always easily identifiable. So it became the librarian’s duty to keep up to date with all the latest enhancements and to maintain regular instructional campaigns, much more so than when one dealt with the print

*Chemical Abstracts*. And while this paper will not discuss every enhancement of each version of SFS, the most important trend over the decade for SFS was the idea of adding additional features that would aid the user in either searching or analyzing the underlying “data” within the documents indexed. So additional features were added that would allow for the analysis of structures and reactions, similar to those available for references. This allowed one to search for a set of similar reactions and find trends such as, what are the most common solvents used for these types of reactions, what is the most common yield to be expected when conducted these sorts of reactions, etc.?

The addition of various limiters to subsequent versions of SFS has also allowed users to further target a search without having to come up with the proper keywords or even narrow a search when keywords would not even be appropriate or possible to use. For example, in SFS, one can search on the roles of compounds, and when one is retrieving all documents related to a particular compound, the user can limit to only retrieve those documents where the compound has been part of a biological or analytical study, or where the compound’s properties are investigated. Again, the trend with some of these features was to use text sparingly to create an initial large answer set and then use the various database features of analysis, limits, etc., to then narrow the answer set to the most appropriate items of interest. These new features allowed for a shift toward a different, yet efficient way to explore topics, substances, authors, etc., especially if the user was unfamiliar with the subject matter. The print was not very efficient for this type of searching since it was indexed in discrete volumes related to a distinct time period. The online version allowed many years to be searched at once and then all relevant results to be gathered and analyzed.

And while structure searching still seemed to be a more unique feature among chemical information databases, additional resources added this feature to their offerings as the decade progressed. For example, the online edition of the 89th edition (2008-2009) of the *CRC Handbook of Chemistry and Physics* introduced structure and substructure searching as a new feature (Roth 2009). While there have been plenty of opportunities for students and faculty to be exposed to structure searching, the ability to conduct a structure search in such a core reference work allowed the concept of structure searching to be introduced to a much wider audience. Even non-chemistry majors often use the *CRC Handbook* for their introductory chemistry courses. In this way, more advanced searching features and capabilities trickled down to a wider audience by becoming available in a larger number of resources. The availability of these additional features or types of searches also made it easier for library users to “graduate” to the more sophisticated databases if they first experienced commonalities between the layout and options available to them after they became familiar with more fundamental resources. So if introduced to structure searching via the *CRC Handbook* at a much earlier level, more users looking at SFS for the first time should feel a little more comfortable initiating a structure search on their own.

### **Improvements to Existing Resources: Platform**

During the last decade, online access to some of the most important chemical information databases was available mainly via client software (*Chemical Abstracts* via SFS and the Beilstein/Gmelin databases via Crossfire Commander) or via CD-ROM (Cambridge Structural Database). This software would often have to be distributed to individuals to be loaded on each of their workstations, networked via some internal environment, or distributed via thin-client software (Culp 2002). This caused chemical information professionals to often add technical

support to their duties and to require further partnerships with internal information technology departments. It was not until the latter half of the decade that these major holdouts began to migrate toward a web interface. The Chemical Abstracts Service introduced SciFinder to the web in 2008, Beilstein/Gmelin first became available on the DiscoveryGate platform in 2004 and is now also available via the Reaxys interface since 2009, and the Cambridge Structural Database for crystallography also introduced a web version in 2009 (Chemical Abstracts Service 2008; Elsevier 2004, 2009; Cambridge Crystallographic Data Centre 2009). The availability of these resources on the web is a much greater advantage now that all users will have access to the latest software and that maintenance of internal distribution networks for client and CD-ROM software can now be minimized. And while sometimes the same features that had previously been available in the client software version might not have immediately and/or fully migrated to the web-based version, the ability to access these resources more easily off-campus, via more operating systems, or on the fly within a classroom or office setting where the software might not normally have been installed, currently off-sets much of this inconvenience.

The web-based delivery of these resources also allows more extensive beta testing because a beta site can be more easily distributed to development partners via a web link and password-protected compared to the old system where software would have to be pushed out and installed before it could be fully tested. Also, the web versions allow for new features to be incrementally released throughout the year (such as quarterly) compared to the previous model where it was usually the norm to release new software updates at most once per year.

Accessing these new resources via web-based interfaces was not without its challenges as more widespread communication and training was needed to make sure all users were aware of

such a drastic change in interface, especially when the client software was no longer going to be supported.

### **Integration with Other Resources**

While the previous sections discussed improvements to existing resources, such as adding additional features or migrating to a different platform for distribution and access, new protocols were developed that facilitated the integration of existing resources, extending the usefulness of a particular resource. Perhaps one of the most important examples of such integration was the adoption of OpenURL linking by various databases. Developed in the late 1990s, OpenURL linking basically allowed databases that previously did not offer full-text documents to attempt to link to the full-text documents via a standardized protocol (Van de Sompel and Beit-Arie 2001). Therefore, nearly any database that adopted the OpenURL protocol effectively became at least a partial full-text database based on access via the particular library's journal subscriptions. Along with implementing the OpenURL protocols, there was again an educational campaign necessary for library users to understand the concept of OpenURL linking. One benefit became the ability to brand the OpenURL links with a custom designed logo or text, allowing the same OpenURL links to appear across all databases where the OpenURL protocol was implemented. This helped reduce confusion among library users and provided some visual consistency across different resources.

Another integration occurred when particular databases offered the use of Persistent Uniform Resource Locators (PURLs) (OCLC 2010). Typically when conducting a search within a database, a results list was generated dynamically and therefore the URL could not be easily saved or bookmarked for future reference. However, particular databases began to offer the ability to create PURLs based on a particular search or even to a specific document within

the database. These PURLs would be a permanent pointer to the search results or document that could be re-initiated at a later time. This allowed for the ability to share search results with others and to place permanent links within course software for students to use. For example, the use of PURLs has also allowed the American Chemical Society to create a PURL for each article on the ACS Publications web site linking directly to the article's corresponding record in the web version of SciFinder (American Chemical Society 2008b). This allows for immediate recall of the particular article in SciFinder and thus a quick entry into finding additional information about a particular author, compound, or topic discussed in the original article. The benefit is also more widespread. While one can often link to other articles by the same author on a particular publisher's web site, one is obviously limited to only those articles published through that particular publisher's portfolio. Linking first into an indexing and abstracting service such as SciFinder opens up one's search across multiple publishers and even non-journal literature.

Finally, there were trends in integrating chemical information resources with other resources found within the research workflow. One of the most crucial integrations in the past decade was between information resources and bibliographic management software. While it was convenient to use bibliographic management software to manage a personal library of relevant citations and then format them into bibliographies or works cited lists as needed, the ability to export citation information directly from the results of a literature search greatly affected the workflow efficiency of many researchers. Search results could also be immediately exported and then reviewed at a later convenience, thus not requiring the need to remain in the native search interface to review results. EndNote 6, released in 2002, for the first time allowed the ability to manage file attachments related to citations (ISI ResearchSoft 2002). This allowed researchers to not only create a personal library of references, but also manage the full-text PDFs

or the supplementary spreadsheet data files as well. While each of the connections between an EndNote citation and a PDF document needed to be made manually, EndNote X2, released in 2008, contained a feature that would attempt to download the full-text PDF automatically, thus saving the researcher even more time, especially if the full-text was not available in the database in which the original citations was retrieved (ISI ResearchSoft 2008).

### **New Resources**

While one major chemical information resource celebrated its 100th anniversary in the 2000s, other influential resources and services first debuted during this decade (Chemical Abstracts Service 2007). One of the most significant was the debut of Knovel in 1999. Knovel partners with other scientific publishers to digitize and enhance their publications. For example, Knovel will take data found within a handbook and transform it into an interactive table that can be searched or it will take equations and graph them so that they can be immediately solved across a wide range of values (Knovel 2010b). By collecting publications on a particular subject area across publishers, Knovel brings together more of a critical mass about a particular topic that is not biased toward any one publisher's portfolio. In particular the Chemistry and Chemical Engineering subject area contains well-known titles such as Hawley's Condensed Chemical Dictionary, Lange's Handbook of Chemistry, Perry's Chemical Engineers' Handbook, and Purification of Laboratory Chemicals (Knovel 2010a). While a subscriber does not own the content (a pitfall that will be discussed later), the power to search across major reference works, including performing numerical property searches can be taught as an important precursor to property searching within a more complicated database such as Reaxys (much like the concept of starting with basic structure searching in the *CRC Handbook* discussed earlier).

While not an entirely new product, another method whereby documents were repackaged or enhanced occurred when the Royal Society of Chemistry launched Project Prospect in 2007 (Royal Society of Chemistry (U.K.) 2010b). Project Prospect enhanced articles have a secondary information layer over the original HTML document. This layer allows the reader to highlight various chemical compounds or terms mentioned in the article and then when those compounds or terms are selected, the reader is shown either the structure or a definition with options of linking to additional articles that discuss that compound or concept. It can be taught to library users that this secondary layer is an efficient tool for either exploring unfamiliar concepts or can be used to retrieve similar articles without conducting an entirely new search. One is simply following a thread of connected ideas or concept by using these additional tools.

Another trend over the past decade was the conversion of major reference works and book series to an online format. While some of these book series are indexed in various databases, one is able to discover much more of their content now than often the entire full-text of a book series, such as all 55 volumes of *Progress in Inorganic Chemistry*, can be searched at once instead of looking at each table of contents or index. Additionally, often these book series are simply cataloged under the main title in a library's online catalog, thus leaving the user without any idea of the large amount of content found within the multiple volumes of a series. Even if one does not subscribe to some of these reference works or books series, still being able to freely search through the content has become a time-saver. Perhaps one of the biggest obstacles to the adoption of regular use of these reference resources by library users is their competition with the likes of Google, Wikipedia, and other freely available web resources. Often these reference works are spread across publishers and are only searchable a title at a time, unlike some freely available web resources that seemingly cover all subjects at once. Perhaps to

counter this, a number of resources, such as Knovel and CRCnetBASE, emphasize cross-searching across multiple reference works and other books, and continually upload newer editions or updates to older works to compete with the continual updates that many free web sites automatically undertake.

### **Information Literacy Standards**

While most of the previous section has stressed the trends among particular resources and their features and how they affected instruction and discovery, the trend toward more specific information literacy standards over the last decade was in some ways a contrast to this tool-oriented perspective. These standards were often developed as a way to emphasize the importance of gaining the knowledge to find, evaluate, and use information independent of source or specific tools. Consequently, there is not as much emphasis on the particular tools to find information, but more on the skills necessary to properly form a search strategy, evaluate results, and ethically use the information one finds. And while information literacy has been discussed since the 1970s, it was not until the 2000s that chemistry specific information literacy standards and guidelines were developed (Zurkowski 1974).

The American Chemical Society's (ACS) Committee on Professional Training (CPT) requires that those undergraduates earning an ACS-approved degree receive training regarding the proper use of the chemical literature (American Chemical Society 2008c). Though the guidelines mentioned only amount to a paragraph in the main document, a Chemical Information Retrieval Supplement was also developed and released in 2008 (American Chemical Society 2008a). This Supplement further details some of the topics that should be taught and mentions skills students should learn before graduating. However, these guidelines are still open to

interpretation and chemical information professionals have looked to other standards as a way to bolster the ACS requirements.

In less than a decade, information literacy standards and guidelines were progressively developed from those necessary for a general academic audience, to those specifically in the sciences, to those in chemistry. First, in 2000, the Association of College and Research Libraries (ACRL) released “Information Literacy Competency Standards for Higher Education” (Association of College and Research Libraries 2000). This established a specific set of standards as well as accompanying performance indicators and outcomes. For the first time a single document could be used by academic librarians to not only discuss specific information literacy outcomes amongst themselves, but also to have a document of support to use to further the advancement of information literacy with non-library faculty and staff at their home institutions. Then in 2006, the Science and Technology Section (STS) of ACRL released more specific standards related to topics relevant to science, engineering and technology (ALA/ACRL/STS Task Force on Information Literacy for Science and Technology 2006). These mirror the 2000 ACRL standards, but the outcomes are further defined for the sciences and related disciplines. Further, in 2007, the Chemistry Division of the Special Libraries Association (SLA) developed “Information Competencies for Chemistry Undergraduates: The Elements of Information Literacy” (SLA Chemistry Division 2007). These guidelines are meant to bridge the ACRL Information Literacy Standards with the ACS CPT requirements. To this end, these guidelines include skills that should be gained, as well as offer specific resources or tools that could be used to impart these skills.

### **Changes in Collection Development Practices**

Over the past decade, the shift from primarily print resources to primarily online only resources caused traditional collection development models and practices to be radically reconsidered. While these changes were primarily influenced by financial considerations and licensing models, the bulk of these changes centered around two areas: the librarian's perception of what the institution is really paying for and the librarian's (and user's) perception of what the institution is really providing access to. These two perceptions may seem similar, but the first is more of an internal perception related closely to acquisitions and licensing, while the second is concerned more with how library users might perceive the outcomes of a librarian's collection development decisions. And while some of these issues were developing in the late 1990s, their full impact was not realized until the 2000s.

### **What Is the Institution Paying for?**

Once publishers and other information providers were freed of the confines of a print work, they began to, and still do, experiment with different delivery and licensing models for their electronic content. This continuous experimentation and realignment of purchase and access models has caused major shifts in collection development attitudes and practice. These shifts can be grouped into five broad areas that forced librarians to change their collection development practices in the last decade: print vs. electronic, purchase vs. lease, new purchase vs. repeat purchase, data vs. platform, and gathering vs. separating.

### **Print vs. Electronic**

Electronic journals (e-journals) existed well before 2000 (Turoff and Hiltz 1982). However, once established, perhaps the first dilemma facing a librarian interested in purchasing an e-journal subscription was dealing with the various pricing models. In the past one could simply subscribe to a print journal for a fixed price, receive the specific journal issues paid for in

a subscription year, and then renew the following year at usually an incrementally higher price. When information providers first began to provide access to their print journals in an electronic format, there was, and still is, no standardized way to charge for the electronic format. At first, some information providers simply provided online access free with a paid print subscription. Others allowed one to pay an additional nominal fee to gain electronic access. As the decade progressed and publishers felt the move away from print, these models were sometimes reversed. In some cases publishers had one fee for a subscription that included the print and online version with no decoupling of the two, while other publishers provided online access with a standard paid subscription and charged an additional fee if a library wanted print copies. This made it difficult for librarians to make collection development decisions because it was often difficult to truly answer, “What is my library paying for, and how would what we pay be affected if we changed formats?” Each publisher is different and therefore a decision, for example, of canceling all print journals across the board might net some savings with certain publishers, provide no savings from others, and actually not be possible with some publishers if there was no decoupling option available.

Another factor in choosing format occurred when some subscription rates began to be based on the size of an institution, creating a tiered pricing structure. While seemingly a fair model relating the number of potential users to an increased subscription cost, simply reporting the size of an institution could skew the fact that a very large liberal arts university might have a very small chemistry department, yet they are paying the same cost for some journals as a land-grant university that has a very robust and active chemistry department. So as some cried foul to this sort of pricing, some publishers have responded with alternative pricing models. Models can now also be based on some other classification describing the level of research conducted at

the subscribing institution (such as Carnegie Classifications) or even on vendor generated criteria (such as previous print subscription holdings, past online usage, presence of particular degree programs, etc.). While these models do help in creating fairer pricing structures, they also require additional involvement from the librarian in order to track down the information needed or to negotiate with the publisher about where the institutions falls within the publisher's tiered pricing structure. With these various models being used by information providers, it is often frustrating not to be able to retrieve straight forward pricing information from a product's web site since pricing can now be unique for each institution. In the traditional sense, a subscription is a subscription, simply is no longer true.

### **Purchase vs. Lease**

As librarians became interested in understanding what exactly it was they were paying for—print, electronic, or print plus electronic—the issue of ownership versus leasing began to emerge as an ever important consideration. While not an exclusive concern of the last decade, as libraries decided to move toward dropping print for a majority of their subscriptions, it became very important to understand what ownership clauses in the license actually entitled a library to in terms of ownership of previously subscribed material (Hawbaker and Wagner 1996). Again, there was no standard practice and for some publishers, one might be entitled to perpetual access to the online version for the years purchased (often for some sort of annual “access fee”) while others might entitle one to nothing or a bundle of CDs or DVDs with the appropriate journal article PDFs loaded onto them.

As collections budgets became squeezed over the decade it created a very difficult situation if one wanted to convert their journals to online only. If budgets continued to shrink and cancellations were necessary, would the institution loose access to the years they had

previously subscribed to? Or should a print copy be held for as long as possible taking the attitude that the publisher could not to be trusted, even if one's license entitled electronic access in perpetuity? But how can a librarian justify purchasing the same content in two formats, especially if savings could be realized by dropping one format or another? And should every institution bear this responsibility given the economic and infrastructure costs necessary to retain large runs of print journals? The response was a series of initiatives to address these concerns. In 1999-2000, Stanford University alpha tested the Lots of Copies Keeps Stuff Safe (LOCKSS) program that has since grown to hundreds of international participating libraries and contains materials from over 400 publishers (LOCKSS 2010). LOCKSS is a program that provides expertise and tools necessary for individual libraries to collect and preserve their e-content. Further, Controlled LOCKSS, or CLOCKSS, was later developed to deal with orphaned content that was no longer under a specific publisher's control. This program is meant to ensure that no published scholarly digital content is lost and if orphaned, would be available to the entire world for free (CLOCKSS 2010). A complementary program, Portico, was introduced in 2002 (Portico 2010). While LOCKSS is a service that a library runs locally in order to preserve only the content they subscribe to, Portico is an off-site service that collects all of the online content from participating publishers in one place and charges participating libraries Annual Archive Support payments in order to support and provide access to orphaned content. With these programs in place, concerns over e-content disappearing have been reduced and therefore the need for all libraries to maintain an archival hard copy has also lessened.

### **New Purchase vs. Repeat Purchase**

As technology helped push publishers forward in providing access to their journals, it also allowed them to look back and provide electronic access to their back runs of journals,

including those no longer being published. While each publisher took a different tact to offering these backfiles to library subscribers, an issue that often arose was how could a library justify paying for something they may have already purchased in the past. The idea of paying money to receive access to the latest journals article as they were published made sense because that was new information, but the idea of purchasing or even leasing electronic access to a seventy year run of a journal that was already sitting on the shelves seemed difficult to justify. So what exactly is a library paying for in this case? The justification came from both sides, library and vendor, that while it might appear that information is being purchased twice, additional costs were incurred by the information provider in terms of digitizing the backfiles, adding metadata, hosting the files on their servers and providing bandwidth to provide access to all users at an institution. Also, the online backfiles acted more like a searchable database instead of individual volumes sitting on a shelf and therefore might have be considered a new “product” for libraries and their users. Unfortunately to outright purchase many of these backfiles, there can be very large one-time costs. While there are often options to lease as well, there is always the need to calculate the rate of pay-off. For example, a different decision might be made if the one-time purchase equals five years of leasing versus twenty-five years of leasing, as well as considering whether ongoing costs for multiple products will continue to add-up and inflate at a greater than manageable rate. In this case the one-time purchase becomes the lesser of two evils in terms of purchasing it now and then being able to forget about additional costs, though some products still require a reduced access or maintenance fee to be paid yearly to the information provider. These types of large one-time purchases often caused librarians to create “wish lists” and line-up for year-end monies that might be available in central funds near the end of each fiscal year.

### **Data vs. Platform**

This issue is related closely with the purchase vs. lease issue, but often involves multiple vendors or information providers.

As discussed earlier, one of the trends in the last decade was for some information providers to repackage material from other publishers. However, because these offerings were often controlled by license agreements between the two parties involved, these offerings could just as easily disappear as they appeared. In this case, librarians had to question what it was they were actually paying for regarding these services. In the case of Knovel, does the library simply pay for the platform that serves up the enhanced data and handbooks, or is the library paying for the data and handbooks themselves? In reality, it is probably both because Knovel needs to develop and maintain their product and Knovel needs to pay the publishers royalties for using their publications.

A prime example of the data versus platform issue occurred in 2003, which involved the *CRC Handbook of Chemistry and Physics*. The *CRC Handbook* was initially available via the Knovel platform and included interactive features. It was also a significant enticement to subscribe to Knovel. However, as CRC Press developed its own online interface, the publisher eventually pulled the *CRC Handbook* from Knovel at the end of 2003. Luckily, Knovel was able to bolster the quality of its Chemistry and Chemical Engineering collection by continuing to add other well-known handbooks, thus minimizing the loss of such a well-known title. Since Knovel does not own the rights to the works from other publishers, their product can be affected if a publisher terminates their agreement with Knovel.

Similar examples involving Knovel still occur. There is sometimes an embargo on new editions of handbooks and they are not released on the Knovel platform until a certain time after the print version has been published. Again, Knovel does not have as much control of

when the publication is released online compared to if they were the actual publisher. Overall the collection of Knovel remains relatively stable, but it is important to realize that what you pay for one year in terms of specific title access might not be the same the next.

Another example in the last decade involved the Beilstein and Gmelin databases, a major source of information on experimental data for organic, inorganic, and organometallic compounds. In March of 2004, MDL, an Elsevier subsidiary, began to offer Beilstein and Gmelin via the web-based DiscoveryGate interface (Elsevier 2004). Any web-based version would eventually become a significant advantage for Mac users since the later version of the Crossfire Commander client software used to access Beilstein and Gmelin would not be supported for Macs (Elsevier 2008). In 2007, Elsevier, through MDL, purchased the data from the Beilstein database from the Beilstein-Institut (Elsevier 2007a). At this point Elsevier owned the data and the interface. However, only several months later, Symyx acquired MDL but Elsevier kept the rights to the Beilstein data. So, Symyx was left providing access to the Beilstein data via the DiscoveryGate interface, but they did not control the data (Elsevier 2007b). As to be expected, Elsevier then began working on its own web delivery platform for the Beilstein and Gmelin data (now called Reaxys) and announced that the CrossFire Commander platform will no longer be supported after December 31, 2010 (Elsevier 2010).

Therefore, in some cases, it is necessary to consider whether one is paying more for a platform versus the data, or both. And even then, through mergers and acquisitions, the data can become decoupled from the platform, leaving much uncertainty.

### **Gathering vs. Separating**

As a consequence of the Data vs. Platform issue, information became available from a variety of vendors, allowing the same data to be accessed from multiple interfaces. This allowed

librarians to make decisions regarding either gathering as many resources under one interface or pairing the best interface with the resource being provided. Initial considerations included ease of use of the interface for library users and the cost of accessing the same resource between different providers. However, as the decade progressed and particular vendors began to also offer full-text documents in aggregate, balances were tipped in their favor compared to another vendor who might offer the same indexing and abstracting resource without linked full-text. Usually there was an extra cost for this, but in terms of instantly connecting the user to the full-text became a clear positive if costs were reasonable.

In terms of acquisitions issues, the idea of gathering as many resources under one provider was appealing because then there were less licenses to deal with, fewer payments to be made to multiple vendors, and the maintenance of web links and other technical issues became easier with fewer base URLs to deal with. The drawback was giving a particular vendor a greater monopoly over an institution's information access, giving rise to the possibility of the vendor raising fees at a higher than usual rate knowing the institution would find it difficult to switch all of the products being offered to new vendors. As always, the relationship is a matter of balance since the institution could also exert some leverage if the institution threatened to pull all business from the vendor.

### **What Is the Institution Really Providing Access to?**

Regardless of how and why the collection development decision is made, the perception of what is actually being offered to the library user is also important, especially from the user's perspective. Regarding the perception of access, major issues over the past decade involved name recognition, critical mass, currency, transparency, and free vs. paid resources. Also,

because these issues involve the library user's perception, opportunities for instruction and education present themselves, similar to the issues discussed in the first section of this article.

### **Name Recognition**

Wikipedia has been called Wikipedia since it began in 2001. Google has been called Google since it was first accessible to the public in 1998. In academia, *Chemical Abstracts* has morphed into SciFinder Scholar (SFS) and now simply SciFinder. The Beilstein and Gmelin databases have gone through a similar evolution from their namesakes to an emphasis on the use of Crossfire Commander, to the most recent manifestation, Reaxys. This tendency to change names, either due to a new interface, new access point, a new company name, etc. can possibly confuse users who know they used to search Science Citation Index, but have no idea that they now need to go to Web of Science to search the same information. Consider if all hardware stores started calling hammers "pounding sticks" and only acknowledged the use of that word? While information providers are building a brand name, the change of name of some of their major products can confuse library users who created a mental tool kit of useful resources to meet their information seeking needs. Unnecessary burden is placed on librarians (and sometimes the vendors marketing staff) to re-educate library users to the simple act of name changes when all involved would much rather be interested in simply learning how to use the information "tool" effectively.

Name changes were sometimes justified, as SFS did initially search the *Chemical Abstracts* data as well as the data from the MEDLINE database, making it more than simply "*Chemical Abstracts* online." However, choosing names drastically different from the core, originating data set could potentially confuse users. Though another possibility could be that the database evolved and the name did not. While SFS has changed to simply SciFinder (dropping

the Scholar moniker for academic use), it has added different types and collections of data since its original inception, making it even more robust than its original intent, yet not changing its name. SFS first only covered *Chemical Abstracts* back to 1967, so it was not a very deep index. But over the years records were incrementally added so that by February 2002, SFS was able to search data back to 1907 (the equivalent of the first volume of *Chemical Abstracts*) and by the end of 2004 even pre-1907 data was added that was not originally indexed in *Chemical Abstracts*. Calculated property data and experimental property data was first attached to chemical compound records in the CAS Registry in 2001 and 2004 respectively, and in 2005, spectra began to be loaded (Chemical Abstracts Service 2010; Wagner 2006). So in this sense, librarians and users must also adapt their conceptions of what the name really means, for examples, as SciFinder can more legitimately be used to search for property data now more so than in 2000.

Another trend in the last decade as journals began to be regularly published online was for some publishers to create “virtual journals” which in effect were a repackaging of individual journal articles from across the publisher’s own portfolio into a new online collection of virtual “issues” centered around a particular topic (American Institute of Physics and the American Physical Society 2010; Royal Society of Chemistry (U.K.) 2010a). Multiple publishers could also contribute to a single virtual journal in a way to draw attention to each of their individual journal titles. The idea of a virtual journal bringing together similar articles across multiple journals was interesting and could save time in searching, but it also ran the risk of alienating library users if they did not have access to all the articles in the virtual journal because their institution did not subscribe to each individual journal title incorporated into the virtual journal. It also threatened to dilute the name recognition of the individual journals themselves.

Regardless of their effectiveness or hampering regarding name recognition, the proliferation of virtual journals has been slow, and some have even stopped “publishing” (Elsevier 2006).

Publishers who had a very large journal portfolio also began to market their online journals more as a collective database instead of a collection of titles. Library users might mention searching “ScienceDirect” or “ACS Pubs” to begin a literature search, not realizing that these are not true indexing and abstracting services, covering multiple publishers, but rather very closed (though sometimes quite large) systems of information access. There is a risk of dilution, whether positive or negative, if researchers simply begin claiming, “My article was published on ScienceDirect.”

### **Critical Mass**

Library users exposed to massive collections of information and data, such as through Google and Wikipedia, appeared to also succumb to a critical mass issue over the last ten years. In the print environment, often multiple resources would be consulted, multiple volumes of indexes were used to narrow a search, and numerous articles were bookmarked for photocopying. However, the idea of finding the path of least resistance, this critical mass of information in one place, caused library users to look for the largest, easiest databases to search that would return the highest number of full-text results. This attitude also filtered down to the universe of paid resources often found at institutions.

While well-known and still heavily used, would a user prefer to search the single title of the *CRC Handbook of Chemistry and Physics* or the *Merck Index* online, or would they prefer to search the plethora of chemical property handbooks on Knovel? The idea of increasing recall instead of increasing precision seems to permeate this attitude—wanting to search many handbooks in the hopes that one has that elusive piece of data instead of searching one individual

title at a time. This also translates into the depth and breadth of information resources. Would a user prefer to search a massive, deep database such as SFS or to search a small, niche database such as METADEX? And when the idea is to lessen the confusion of library users from the myriad of databases available to them, the preference would seem to lean toward the resources that collect the most information in one place. However, it is critical that library users also understand that these niche or individual resources often have some unique aspect that makes them wholly appropriate in certain situations. So while the Cambridge Structural Database indexes many journals also indexed by SFS and would seem to be duplicative, it also provides data on individual crystal structures that is totally absent in SFS. Therefore, it is still important for library users to understand what resources are being offered and the pros and cons of each information resource.

### **Currency**

Another issue facing library users is interpreting the offerings of their libraries and how it relates to the currency of the information resources available to them. In terms of journals, this is relatively easy since many journal homepages default to the latest issue or even highlight articles that have been published online ahead of the print publication or in an official online issue. This issue is more relevant to indexing and abstracting services and to online reference works.

As article databases have developed online, they have also developed the ability to index articles in a more rapid manner. Databases such as SFS and PubMed index articles daily, even if only offering a basic record until more extensive indexing can be accomplished. On the other hand, Beilstein and Gmelin via Crossfire are only updated quarterly. In terms of searching the latest research, this is something that library users need to be properly made aware of since not

all databases are equally current. When users are presented with blank search boxes in each interface, it is more difficult to figure out the currency of a database compared to looking for the copyright date or within the preface of the work to find out how old the information inside might be. Google and other internet search engines have given library users a false sense that other online searchable databases are also (seemingly) instantly up-to-date.

Finally, online reference works present special problems related to the perception of the currency of the material. While again, Wikipedia may be viewed as having the ability to instantly be updated, online reference works have various publishing models depending on the publisher. The notion that an online reference is a hybrid between a discreet single edition much like the print and a growing reference work able to be updated like Wikipedia makes it problematic when presented in the online environment. For example, the Kirk-Othmer Encyclopedia of Chemical Technology published by Wiley can be purchased for a one-time fee and users have access to a particular edition of the encyclopedia online. However, an institution can also subscribe to Kirk-Othmer for an ongoing reduced fee and have continual access to updated articles as they are published online. The drawback is that if an institution stops subscribing to this version, it loses all access. Additionally, once a critical mass of updates have been published, it might be sufficient to indicate a new edition of the online product, allowing for the one-time purchase of this new edition. So, while there are two distinct models—a one-time purchase of a static work or the serial subscription to an evolving work—to the user, there is no difference in the online experience. The interface at Wiley's web site looks exactly the same for either version of Kirk-Othmer. While the individual encyclopedia articles do indicate when they were last updated, two different users at two different institutions might be searching the same

title yet much different content. Again, library users do not seem to be keenly aware of these issues and how it might affect their searching.

In either of these cases involving currency, it has become more difficult in the online environment to determine how current and dynamic the content is that is hidden behind a search interface.

### **Transparency**

The ability of library users to discover how current a database is also relates to the issues of transparency in the online environment. Transparency is not only what a particular resource is, but in terms of what is actually being offered by a library user's particular institution versus what is more freely available on the web. While it is every institutions' hope to provide a seamless experience, this is very difficult to achieve with 100% accuracy.

Libraries have increasingly steered their users to an online portal or gateway in order to go through a proxy or similar service to access information resources that their institutions has licensed for them. However, one problem with this is that often, the proxy is only needed when the user is outside the institutions network and needs to authenticate themselves as being a valid users of the institution's network. Therefore a user on a university campus can access many resources by typing in a specific URL or by going through a bookmark without having to go through the proxy. However, when this same user is off-campus, typing in a URL or using that same bookmark most likely will not work. So behavior needs to be changed compared to on-campus and off-campus behavior, or the same behavior necessary off-campus should be practiced on-campus in order to reconcile these two methods of access. This can also confuse library users because off-campus they can get to free resources without going through a proxy, so

the thought of seamlessly going from one resource to another or one site to another is interrupted by the need for the user to authenticate themselves in certain instances, but not for others.

In the last decade many information providers have now allowed subscribing institutions to include a brand or logo on the publisher's web site in order to reinforce the nature of the relationship between the subscribing institution and the vendor. This works as a marketing tool to raise awareness of the fact that certain information resources are paid for compared to the assumption that the information was freely available.

The implementation of the OpenURL protocol allowed institutions to link their full-text holdings across various databases and other services. This helped greatly in seamlessly linking the user from a specific citation to the desired full-text document. Since OpenURL is an accepted standard, the ability to use it with free resources such as Google Scholar, has blurred the boundary between free and paid resources. So even library users searching Google Scholar have a chance to retrieve full-text documents to which their institution subscribes.

### **Free vs. Paid Resources**

Finally, as library users are exposed to various resources at their institution, another hurdle facing them is when they switch institutions (academic, corporate, not-for-profit, etc.), and coming to the realization that what they have access to is different. Librarians walk a fine line in terms of educating library users on the costs of certain information resources versus teaching free resources. As the latter half of the past decade created more stressful financial situations for many institutions, librarians even began to question whether certain free resources could replace paid information resources.

The information literacy standards discussed earlier get to the heart of this issue. Instead of only teaching library users one tool after another, it is also important to teach them about

information in general—how it is structured in their discipline; how to identify information needs and the type of information that could fulfill these needs; how to evaluate various resources and information found in these resources; and how to place information within various social, legal and economic contexts. One might argue whether it is worthwhile for a chemistry undergraduate student to learn SFS when they might go to a small start-up and never have access to SFS again? Why should they learn SFS when they will have to rely on free web sites or other internal databases?

One of the first lessons that should be taught is that, for better or worse, all information is not free. It should be understood that information resources should be placed in a continuum in terms of content, search capabilities, currency, etc., and that when library users are placed in a new or different situation, they realize they have the more abstract skills required to find the information they seek by any means available to them.

## **Conclusion**

The rapid development of online resources in the last ten years has caused chemical information professionals in academia to adapt their instruction techniques and even the focus of what they are actually teaching. A balance must be made between a straightforward teaching of specific tools and features versus providing users with information seeking and processing skills they can use regardless of the resources they are confronted with.

The last decade also forced chemical information professionals to rethink traditional methods of collection development, paying particular attention to what it was they were actually paying for in the less than straightforward online environment. These decisions and the online resources offered also forced library users to reevaluate what it was they were actually being provided access to, often with difficulty without a librarian's intervention (whether direct or

indirect). Search interfaces and capabilities will continue to change, perhaps at an even more accelerated rate. Library users are continuing to make assumptions that librarians should help to dispel or at least dissuade.

The next decade will most certainly provide additional challenges, however some of the issues of the past decade are still with us and will continue to present challenges for the foreseeable future.

## REFERENCES

ALA/ACRL/STS Task Force on Information Literacy for Science and Technology. 2006.

“Information Literacy Standards for Science and Engineering/Technology.” Association of College and Research Libraries.

<http://www.ala.org/ala/mgrps/divs/acrl/standards/infolitscitech.cfm> (accessed May 25, 2010).

American Chemical Society. 2008a. “Chemical Information Retrieval.”

[http://portal.acs.org/portal/PublicWebSite/about/governance/committees/training/acsapproved/degreeprogram/CTP\\_005584](http://portal.acs.org/portal/PublicWebSite/about/governance/committees/training/acsapproved/degreeprogram/CTP_005584) (accessed May 25, 2010).

———. 2008b. “New Web Technology Transforms American Chemical Society’s Online Journals.”

[http://portal.acs.org/portal/PublicWebSite/pressroom/newsreleases/WPCP\\_011225](http://portal.acs.org/portal/PublicWebSite/pressroom/newsreleases/WPCP_011225) (accessed May 25, 2010).

———. 2008c. “Undergraduate Professional Education in Chemistry: ACS Guidelines and Evaluation Procedures for Bachelor's Degree Programs.”

[http://portal.acs.org/portal/PublicWebSite/about/governance/committees/training/acsapproved/degreeprogram/WPCP\\_008491](http://portal.acs.org/portal/PublicWebSite/about/governance/committees/training/acsapproved/degreeprogram/WPCP_008491) (accessed May 25, 2010).

American Institute of Physics and the American Physical Society. 2010. “Virtual Journals in Science and Technology.” <http://www.virtualjournals.org/vjs/> (accessed May 25, 2010).

Association of College and Research Libraries. 2000. “Information Literacy Competency Standards for Higher Education.”

<http://www.ala.org/ala/mgrps/divs/acrl/standards/informationliteracycompetency.cfm> (accessed May 25, 2010).

- Bolek, Ann D. 2000. SciFinder Scholar. *Issues in Science and Technology Librarianship*  
<http://www.istl.org/00-fall/databases.html>.
- Cambridge Crystallographic Data Centre. 2009. "WebCSD v1.0 released."  
[http://www.ccdc.cam.ac.uk/news/archive\\_2009/](http://www.ccdc.cam.ac.uk/news/archive_2009/) (accessed May 25, 2010).
- Chemical Abstracts Service. 2007. "CAS 100th Anniversary." American Chemical Society.  
<http://www.cas.org/aboutcas/cas100/index.html> (accessed May 25, 2010).
- . 2008. "New SciFinder Available Via the Web." American Chemical Society.  
<http://www.cas.org/newsevents/releases/scifinderweb040708.html> (accessed May 25, 2010).
- . 2010. "SciFinder Solutions Archives." American Chemical Society.  
<http://www.cas.org/support/scifi/sfsolutions/archives/contentarc.html> (accessed May 25, 2010).
- CLOCKSS. 2010. "Home - CLOCKSS." Stanford University.  
<http://www.clockss.org/clockss/Home> (accessed May 25, 2010).
- Culp, F. Bartow. 2002. Streamlining Electronic Access with Windows Terminal Server. *Library Hi Tech* 20 (3): 302-307.
- Elsevier. 2004. "MDL® Offers DiscoveryGate Research Platform to Academia."  
[http://www.elsevier.com/wps/find/authored\\_newsitem.cws\\_home/companynews05\\_0010](http://www.elsevier.com/wps/find/authored_newsitem.cws_home/companynews05_0010)  
9 (accessed May 25, 2010).
- . 2006. "Proteomics Virtual Journal."  
[http://www.elsevier.com/wps/find/P04.cws\\_home/vj\\_proteomics\\_main](http://www.elsevier.com/wps/find/P04.cws_home/vj_proteomics_main) (accessed May 25, 2010).

- . 2007a. “Elsevier Acquires the Beilstein Database, World's Leading Archive of Chemical Data.”  
[http://www.elsevier.com/wps/find/authored\\_newsitem.cws\\_home/companynews05\\_00621](http://www.elsevier.com/wps/find/authored_newsitem.cws_home/companynews05_00621) (accessed May 25, 2010).
- . 2007b. “Symyx Technologies to Acquire MDL Information Systems from Elsevier.”  
[http://www.elsevier.com/framework\\_aboutus/pdfs/Symyx\\_to\\_Acquire\\_MDL\\_from\\_Elsevier\\_FINAL\\_081007.pdf](http://www.elsevier.com/framework_aboutus/pdfs/Symyx_to_Acquire_MDL_from_Elsevier_FINAL_081007.pdf) (accessed May 25, 2010).
- . 2008. “CrossFire Commander: System Requirements.”  
<http://info.crossfirecommander.com/requirements.asp> (accessed May 25, 2010).
- . 2009. “New ‘Reaxys’ Solution Improves Workflow Quality and Efficiency for Chemists.”  
[http://www.elsevier.com/wps/find/authored\\_newsitem.cws\\_home/companynews05\\_01114](http://www.elsevier.com/wps/find/authored_newsitem.cws_home/companynews05_01114) (accessed May 25, 2010).
- . 2010. “Crossfire Support and Training.” <http://info.crossfiredatabases.com/support.html> (accessed May 25, 2010).
- Hawbaker, A. Craig, and Cynthia K. Wagner. 1996. Periodical ownership versus fulltext online access: a cost-benefit analysis. *The Journal of Academic Librarianship* 22: 105-9.
- ISI ResearchSoft. 2002. “ISI ResearchSoft Today Ships EndNote® Version 6 For Windows®.”  
<http://www.endnote.com/pr-en6.asp> (accessed May 25, 2010).
- . 2008. “The Scientific Business of Thomson Reuters Releases EndNote X2 for Windows.” <http://www.endnote.com/pr-enX2win.asp> (accessed May 25, 2010).

- Knovel. 2010a. "Knovel - Browse Chemistry and Chemical Engineering - General References."  
<http://www.knovel.com/web/portal/browse/subject/50/topic/145/mysub/0> (accessed May 25, 2010).
- . 2010b. "Knovel - Technical Engineering Reference Information."  
<http://why.knovel.com/> (accessed May 25, 2010).
- LOCKSS. 2010. "Home - LOCKSS." Stanford University.  
<http://lockss.stanford.edu/lockss/Home> (accessed May 25, 2010).
- OCLC. 2010. "PURL Home Page." <http://purl.org/docs/index.html> (accessed May 25, 2010).
- Portico. 2010. "Our Organization - Portico." ITHAKA. <http://www.portico.org/digital-preservation/about-us/our-organization> (accessed May 25, 2010).
- Roth, Dana. 2009. "Handbook of Chemistry and Physics - 89th edition (2008-2009)." CHMINF-L Archives. <https://listserv.indiana.edu/cgi-bin/wa-iub.exe?A2=ind0902&L=CHMINF-L&P=R1986> (accessed May 25, 2010).
- Royal Society of Chemistry (U.K.). 2010a. "Highlights in Chemical Science"  
<http://www.rsc.org/Publishing/ChemScience/index.asp> (accessed May 25, 2010).
- . 2010b. "Project Prospect."  
<http://www.rsc.org/Publishing/Journals/ProjectProspect/index.asp> (accessed May 25, 2010).
- SLA Chemistry Division, Ad Hoc Committee on Information Literacy. 2007.  
"Information Competencies for Chemistry Undergraduates: The Elements of Information Literacy." Special Libraries Association. <http://units.sla.org/division/dche/il/index.htm> (accessed May 25, 2010).

- Somerville, Arleen N. 1998. SciFinder Scholar (by Chemical Abstracts Service). *Journal of Chemical Education* 75 (8): 959, 975-976.
- Turoff, Murray, and Starr Roxanne Hiltz. 1982. The Electronic Journal: A Progress Report. *Journal of the American Society for Information Science* 33 (4): 195-202.
- Van de Sompel, Herbert, and Oren Beit-Arie. 2001. Open Linking in the Scholarly Information Environment Using the OpenURL Framework. *D-Lib Magazine* 7 (3), <http://dx.doi.org/10.1045/march2001-vandesompel> (accessed May 25, 2010).
- Wagner, A. Ben. 2006. SciFinder Scholar 2006: An Empirical Analysis of Research Topic Query Processing. *Journal of Chemical Information and Modeling* 46 (2): 767-774.
- Zurkowski, Paul G. 1974. "The Information Service Environment Relationships and Priorities." Education Resources Information Center. [http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content\\_storage\\_01/0000019b/80/36/a8/87.pdf](http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/36/a8/87.pdf) (accessed May 25, 2010).