

Don's Conference Notes

by **Donald T. Hawkins** (Conference Blogger and Editor) <dthawkins@verizon.net>

Video as a Scholarly Content Format: An NFAIS Virtual Seminar

We live in a virtual world in which video is becoming increasingly prominent. An **NFAIS Virtual Seminar** on May 14, 2014 entitled "*The Emergence and Rise of Video as a Scholarly Content Format*" examined the growing importance of video in academic institutions and how they are responding to the demand for it. Following their presentations, each speaker was asked to identify the major issues still needing attention.

Stephen Rhind-Tutt, CEO of **Alexander Street Press** (ASP, <http://alexanderstreet.com>), led off with an excellent overview of the current role of video in the scholarly landscape. ASP was one of the first vendors of commercial streaming video content collections for libraries and has been selling video since 2006, serving hundreds of media companies, film makers, and producers, as well as millions of faculty and students at more than 30,000 institutions. **Rhind-Tutt** said that a major change has occurred in the past six to seven years because no dedicated devices are now required to record or view video: the old "media cart" that was wheeled into classrooms has gone the way of the dinosaurs. Today's students are well versed in video, and many of them have been using it for almost all of their academic life. They watch video in classrooms and on the Web; they record it on their smartphones, capturing lectures, filming experiments, and conducting interviews; and they use Skype to connect with their peers.

Much of today's video usage is via YouTube. One in seven people use YouTube regularly, and 100 hours of video are uploaded every minute. (See <http://youtube.com/yt/press/statistics> for more fascinating statistics on its usage.) YouTube now has twice the popularity of Wikipedia, and video currently accounts for 57% of network traffic, primarily because of the enormous rise of media-enabled tablet computers.

Because video occupies such a large role in the marketplace, it is important that publishers and content providers embrace it. It is no longer trapped and inaccessible on old media, and according to a report on a national survey that was presented at the **2013 Charleston Conference** (slides available at <http://www.slideshare.net/Charleston-Conference/streaming-video-in-academic-libraries-preliminary-results-from-a-national-survey>), 70% of academic institutions now stream video, and those that don't are planning to do so within the next three years. Collections of video content exist at many universities; about 60% of them are using vendor-created portals because it is easy to subscribe to them. The top two leading vendors in this market are **ASP** and **Films on Demand** (<http://ffh.films.com/digitallanding.aspx>). According to a 2009 report, major uses of video by faculty members are as shown here.



Uses of video extend across many disciplines, such as film, dance, education, anthropology, news, and history. It provides new learning methods; for example, students studying dance can not only read about the moves of performers, but they can also see them. Videos of significant moments in history can portray not only what happened but also the background and emotion, thus taking understanding of the event to a different level.

Against the Grain / February 2015

Rhind-Tutt concluded with a list of nine issues necessary to make video useful in academic environments:

1. **Digitization.** Much material is still stuck on old media, and processes need to be developed to get it online.
2. **Quality.** Material must be complete and accurate. It is all too easy for editing processes to make subtle and significant changes in the content.
3. **Searchability.** Detailed search is still not available for videos — most of them can only be cited by their title. Much of their latent value is therefore lost.
4. **Speed of comprehension.** Transcripts significantly lower comprehension times; for example, although 30 minutes of video take 30 minutes to view, a transcript of those 30 minutes is 12 pages long and can be read in depth in 6 minutes or scanned in 2 minutes.
5. **Analysis and annotation** will allow users to jump to different parts of a video or isolate sections for further study.
6. **Library integration.** There is a large demand for database searching and cataloging tools to be added to videos.
7. **Sharing.** Video clips can be e-mailed, embedded in course materials, cited, and saved.
8. **Tablets and mobile access.** Tablets are excellent for video access.
9. **Preservation** is one of the most important issues. Many valuable items have been lost, including 90% of silent films produced before 1929 and half of the films with sound produced before 1950.

The technical barriers to widespread adoption of video have been largely surmounted, and new ways of using video are emerging rapidly. Tomorrow's students and scholars will demand video and will expect it to be a normal part of their curricula.

According to **Rhind-Tutt**, the top issues still needing attention are:

- **Transcripts** which will promote searching and aid discoverability,
- **Metadata** (today's MARC records are still the best we have),
- **Licenses and permissions** (today's situation is similar to where books and journal were in past years), and
- **Demand across all disciplines.** There is much potential for growth in many disciplines, where the ability to see is very important.

Deidre Woods, Executive Director, Open Learning Initiative at the **University of Pennsylvania**, discussed the role of video in changing academic environments. She noted that we are in a period of major disruption, with the emergence of non-traditional students and MOOCs. Students with high expectations of technology arrive at universities well able to use online resources and expect to learn online. **Woods** echoed **Rhind-Tutt's** observation that YouTube has been a major force for change. It is used every day as a reference for something, and even videos created with very low budgets can have significant meaning.

Many sources of institutional support for using video exist, such as libraries, information technology departments, and instructional design materials. Classes are now being "flipped," in which students watch videos of lectures ahead of time and then use their class time to work problem sets in small groups. One advantage of flipped classes is that students can view the lecture material in depth and can rewind and review it to gain a better understanding; however, experience has shown that creating video materials for flipped classes requires a significant amount of time and effort.

Woods concluded with a view of what works with video, what doesn't work, and what's next:

What works:

- Faculty interested in experimentation who are willing to take risks,
- An institutional mandate and support (which is critical),
- Videos in small chunks (about 4 to 6 minutes): students generally will not watch more (most lectures do not need 50-minute periods to get their point across!), and
- Video with content close to class assignments (if it's relevant to the exams, students will readily watch it).

What does not work:

- Poor quality audio (poor video is marginally acceptable but not poor audio),
- Material originally created for another purpose, and

continued on page 80

<<http://www.against-the-grain.com>> 79

- Material not appropriate for today's devices such as tablets and smartphones (content must be delivered where the users are; much of the world does not use PCs any more).

What's next:

- Courseware and adaptive learning (of which video is a significant part),
- Changes to classrooms (what should be done with large lecture halls in an environment of flipped classes?),
- Research on emerging standards and best practices,
- Videos for the humanities (there are many videos used in business, scientific, medicine, and engineering courses, but fewer in the humanities), and
- Experimentation and innovation.

Woods identified these issues still needing attention:

- Global bandwidth for remote dissemination of information,
- Standards for what is acceptable and effective, and
- Getting past a fear of irrelevance.

Michael Stoller, Director, Collections and Research Services, **NYU Libraries**, discussed how today's libraries are interacting with video (**NYU** has been acquiring video content for 30 years). He said that libraries' needs are simple: acquire and store video content, make it accessible to students and researchers in the ways they need it, and preserve it for future generations. In the past, acquisition was easy, and the major adjustment was the change from videotapes to DVDs, but now videos are online. Services such as YouTube and iTunes are well attuned to the consumer market and credit card purchases, but libraries generally buy materials with purchase orders, which is difficult to do online. Much video content is in formats other than those used by services such as **Alexander Street Press**. The library world is still waiting for a vendor that can work with them to acquire those materials.

Storing videos was easy when they were on tapes or DVDs because they could be stored like books and did not take up any more space than a book. But when videos went online, the file sizes became huge — up to 100 gigabytes per hour of video. Thus, **NYU's** collection of commercial videos on tapes or DVDs would require up to 8 petabytes (8,000,000,000,000,000 bytes) of storage. Making all that video accessible to users is not easy, particularly if it must be streamed to locations on the other side of the world. (**NYU** has campuses in Abu Dhabi and Shanghai, and watching a movie streamed there is painfully slow.)

Preservation of video content is a major concern. Many videos only exist in obsolete formats; **NYU's** collection of 65,000 tapes is obsolete because industrial-quality players are no longer available. One might think that they could just be copied to DVDs, but DVDs are likely to become obsolete. And streaming the material to classrooms is not possible because of copyright restrictions; today's copyright law permits usage only in the library. Despite these challenges, video has become our most common form of expression because we live in a visual culture. It has become an essential part of classroom instruction and research — 35% of all **NYU** classes actively use video. Libraries need to deal with video content because they are memory organizations; it's their job!

Stoller said that the main issue needing attention is not technological but intellectual property law and restrictions on usage, particularly for commercially purchased video.

Rick Gilmore, Associate Director and Co-Principal Investigator, discussed **The Databrary Project** (<http://databrary.org>), an open-data library for developmental science where researchers can share video, audio, and related metadata. The project is governed by a community of advisors and is supported by **NSF**, **NIH**, **NYU**, and **Penn State University**. (**NYU** has committed to store data from **Databrary** indefinitely.) **Gilmore** said that **Databrary** concentrates on video because it is uniquely rich, transparent, readily repurposed, and easily captured. For example, he is working with head cameras mounted on small children to capture what they are seeing and help us understand their visual experiences.

A major challenge encountered by the project is the need to preserve privacy and confidentiality. Because faces on videos are identifiable, researchers are required to undergo ethics training and be governed by local ethics boards. Access to project data is restricted to only authorized users, and there is a tie between users and their institutions. Data can be shared only between institutions that have signed a use agreement.

Databrary researchers have developed an open source viewer, **Datavyu** (<http://datavyu.org>), which runs on both Windows and Mac platforms and incorporates video coding and visualization tools. **Databrary's** vision of the future includes partnerships with professional societies in which materials from journals and conferences can be stored and shared, as well as partnerships with publishers to share published manuscripts and supplemental data.

Jim Ulsh, Director, Product Development, **ProQuest**, reviewed some of the infrastructure considerations surrounding video content. He noted that many people are well aware of **ProQuest's** textual content but do not realize that it also has thousands of hours of video content from hundreds of sources. Extensive use is made of cloud-based systems for content delivery, storage, processing, internal processes, and marketing. Scalability is the major reason for using the cloud, but other considerations include delivery optimization, support, availability, costs, and ease of use. Player support is also a significant consideration; newer video players have resolved many platform issues.

It is important to take inventory of content and know what you have before starting a video project because video means various things to people. Formats, conditions, and volumes are also significant. A major consideration is rights issues that will determine what can be distributed or streamed and limitations such as audiences, geography, etc. Demands to remove content from a database will be inevitable, so a takedown process should be established before distribution begins.

Discoverability is a key issue for audiences. How will users search for and find the content they want? How will they use it? Metadata is important; some is better than none. **Ulsh** recommended not sanitizing the metadata because many fields may be unique to the data. Users will expect to be able to access video content on multiple devices, search, browse, and find related content. Transcriptions are highly useful and can be done by machine or manually. Automated transcriptions are cheaper and faster than manual ones, but they are less accurate and lack structure.

Betsy Bollar, Director, **JBJS** Product Line, **Journal of Bone and Joint Surgery, Inc. (JBJS)** began the final session of the Webinar with a description of how **JBJS** is using video. **JBJS** is the leading journal in the orthopedics field and is currently celebrating its 125th year of publication. Its product line includes several related journals, one of which, *Essential Surgical Techniques (EST)*, is devoted to orthopedic surgery. A survey of *EST* readership revealed that 70% of its audience was accessing Websites to watch videos of operations, 97% of the viewers said that the content of greatest interest to them was surgical techniques, and half of them said their preferred video length was 10-12 minutes. So a project was begun to include videos in *EST* articles. The first author-submitted video was added in 2004, and from then until August 2013, 101 videos were added to the journal. Until then, viewers had to navigate away from the article to the video file, which was an inconvenience; however, in August 2013, the capability to add "inline" videos to articles became available. Since then, 93 new videos have been added, and the journal's Table of Contents indicates which articles contain videos. All new *EST* articles now contain at least one video.

A new Deputy Editor for Video position was created, and that Editor reviews all submitted videos using these criteria:

Technical and Production:

- Video clear?
- Camera steady and smooth?
- Can the anatomy be appropriately visualized?
- Lighting acceptable?
- Field of view and angle of view clear?
- Appropriate captions?

Content:

- Video relevant?
- Video accurate?
- Does it meet its stated goals?
- Is the video factual?
- Are the steps shown appropriate or are important sequences missing?

Because *EST* articles discuss surgical procedures, individual short clips 15 to 30 seconds long showing each step in a procedure are preferred over a longer video showing the entire operation.

Brightcove (<http://www.brightcove.com>) was selected as the video storage and streaming platform. **JBJS** developers used its API to create

continued on page 81

an interface allowing authors to upload their videos directly and send the URL to the Editorial Manager. This system has been well received by the authors who found that it significantly enhanced and streamlined the upload process.

Using the video expertise gained in the past year, a video-based certification course with 15 hours of content and offering CME credits has been developed. **Bollar** said that the most urgent needs for **JBJS** are to increase the amount of video on the site and help authors become more familiar with the technology of video editing. She also said that DOIs for individual figures and videos will be created in the future which will make it easier to cite them.

Jane Hannon, Acquisitions Editor, *Journal of Visualized Experiments* (*JoVE*, <http://www.jove.com>), followed **Bollar** and said that *JoVE* is a completely video-based methods journal. She noted that scientific publication has not kept up with advances in the research process, so it is often difficult to replicate research results. In fact, two studies recently published in *Nature Drug Discovery* found that fewer than 25% of the results from a number of articles could be replicated. **Hannon** suggested that if one could see how the experiments were done instead of having to rely only on a textual description, reproducibility would be significantly enhanced.

JoVE publishes articles from many disciplines, each of them containing video clips of experimental methods. Even though the articles are video-based, they follow the same structure as a traditional research article, with an abstract, introduction, methods, results, and conclusions. *JoVE* does not use author-submitted videos; once an article has been accepted for publication, a professional videographer is sent to the author's laboratory to shoot the video, thus freeing authors from the need to worry about the technical details of producing an acceptable video. Most of *JoVE*'s videos are 10 to 12 minutes long. **Hannon** said using professional videographers provides substantially more control over the videos and results in higher quality, so *JoVE* plans to continue this process. The costs are offset by library subscriptions, and authors are asked to contribute as well.

Once the video has been created and published, the author receives a copy of it and is encouraged to share it among peers and colleagues, which increases its exposure by being cited. As *JoVE* says, "Get sighted by getting cited!" *JoVE* articles are indexed in a number of traditional databases, such as PubMed, SciFinder, Chemical Abstracts, and Medline, and they are cited just like articles from any other journal. *JoVE* currently publishes about 80 videos per month.

JoVE has been well received in the market; over 675 institutions are now subscribers, and the Website receives over 400,000 visitors per month, 80% of whom are affiliated with academic institutions, some of which are shown in the graphic at the top right.

Users like *JoVE*, both for reading and for learning about new techniques, as shown by these two typical quotes from case studies:

- From an author: "The video format conveys complicated methods significantly better than text alone and helped validate our novel results."



- From a reader: "Once you can see how someone else does something, or see how someone does something differently than you, it makes your work better."

One of *JoVE*'s readers reported that she saved over \$15,000 in the costs of chemicals, wages, and travel by watching videos published in *JoVE*, as well as a significant amount of time (30 days to learn new experiments without the videos vs. 5 days with them).

Hannon said that the most often requested enhancement to *JoVE* is a mobile application so that readers do not need to be tied to their PCs but can watch the videos on the fly as they conduct experiments, show them to their colleagues, etc.

The presentations were followed by a question period, in which the presenters were asked for their opinions on what people need to know when they are considering a video project or when they are just beginning one. The responses fell into a number of similar themes:

- **Rights.** Permissions to create videos and licenses to use and distribute them.
- **Formats.** How files are labeled and organized; and where they are stored.
- **Metadata.** What metadata is associated with the videos and how they are cited.

Additional observations included:

- Not everything requires video; consider if the content would really be enhanced by it. Start by looking at use cases.
- Consider the costs. Videos don't just materialize from iPhones! They are expensive; make sure you will serve the user appropriately if you produce one.

The program for the Webinar is available at <http://nfais.org/event?eventId=547>, and access to an archival version can be purchased by contacting **NFAIS** at <http://www.nfais.org>.

The Next Generation of Discovery Services: Where Are They Headed? — An NFAIS Workshop

This workshop, held on June 20, 2014, was a follow-up of previous **NFAIS Workshops** on various aspects of discovery services. It attracted an audience of about 80, 30 onsite and 50 virtually, and brought together players from various parts of the industry: librarians, publishers, and content providers.

Discovery Service Overview

Judy Luther, President, **Informed Strategies**, led off with a review of the discovery landscape and noted that the field has been marked by unexpected turns and strategies. Libraries have long wanted to provide a unified Google-like way for their users to seamlessly access all of their content. Such an approach benefits not only the users, but also librarians who see increased use of their resources, and content providers who gain greater visibility of their content. Today's four major discovery services are:

- **EBSCO** Discovery Service (EDS), which is often used as a starting point for research,
- **Primo** from the **ExLibris Group**, offering a seamless integration with a library's other services,
- **ProQuest's** Summon Service, which has recently added a capability to automatically update a library's eBook holdings, and

- **OCLC's** WorldCat, which has entered into a partnership with **NBCLearn** (<http://www.nbclearn.com>) to provide access to primary source videos and is integrated with teaching tools.

According to **Luther**, a major challenge for discovery services is their impact on the visibility of content, which is usually determined algorithmically. The issue really comes down to whose content appears on the first page of results — many users never get farther (which is a common issue with all searching systems, including Google). Other challenges are the need for integration with other tools and services — the back end for delivery and the front for coverage — and a diminished use of advanced features of the content such as specialized thesauri and search strategies.

Content neutrality in a highly competitive market is a major issue for Web-scale discovery systems because it is difficult to avoid bias in displaying search results.¹ There are also complex issues to be addressed in negotiating for access to databases, such as what metadata will accompany the results and whether the content has been prepared for use in a discovery service. Despite these issues, discovery services offer significant advantages to users because they permit searches across a wide range of content, including that

continued on page 82

which is not held in a local library, expanded access to materials used in research projects, ease of use, and convenience.

End users are generally not knowledgeable about resources, and librarians may be reluctant to accept the convenience of discovery services because they feel they can search better than their users. There must therefore be a balance between librarians' attitudes and the need to let users do their own searches when they wish to. The end goal is to put content into users' hands, and studies have repeatedly shown that a large fraction of users begin their searches with a search engine. Faced with this reality, the **University of Utrecht** decided to abandon its discovery tool and focus on delivery.² Libraries understand the value of the resources they are paying for and want to give their users the best possible solution to their needs. The discovery tools have done a very good job at providing a robust platform. Google Scholar is one of their major competitors, but it is not widely trusted because it could disappear at any time. It also suffers from problems of incomplete and inconsistent coverage; therefore, many libraries do not want to move away from discovery services. **Luther** concluded that libraries need to always consider the value that users place on convenience, and they must find ways to ensure that their content is well presented to users.

Discovery Services in Libraries

Scott Anderson, Information Systems Librarian, **Millersville University**, described the use of EDS at his institution (it was launched there in 2010) and noted that most content is used by faculty working with students studying in specific disciplines, but not all of the necessary resources are included in discovery services. About 30% of the searches done at **Millersville** on **EBSCO** databases use EDS; the remainder are on a single database. **Anderson** studied the use of *Biological Abstracts* (BA), which is available as a standalone database as well as on EDS and found that over the last two years, EDS has improved at driving users to BA content. His study concluded that it is good to know what sources a discovery service is driving users to so that a case can be made for purchasing them. Usage of discovery services also causes adjustments with content providers as they endeavor to get more appropriate content to satisfy users' needs.

Millersville's implementation of EDS is configured to show the universe of reasonable resources, not just what the library has. It is not limited to full text or library holdings. The traditional OPAC search box has been replaced by one based on the discovery service, thus putting print and electronic resources on an equal footing (most OPACs tend to emphasize print). The discovery tool handles the catalog, eliminating the need to load MARC records. It provides traditional catalog functions and also drives unmediated purchases of resources such as eBooks. In the future, **Anderson** would like to see improved analytics and coordination with local reporting tools, real-time reports (every 4 hours) of what is occurring in the system and what headings are being searched, and the addition of all the content into the discovery space, including the metadata.

Discovery services are not only applicable to traditional library resources, but they also provide access to "ancillary" services such as course materials, institutional repository materials, research networks, etc. **Franny Lee**, Co-Founder and VP of Business Development at **SIPX** (formerly **Stanford Intellectual Property Exchange**), described the impact of discovery services on ancillary services. **SIPX** (<http://www.sipx.com>) is a Web service for managing and sharing digital course materials in higher education. It interfaces with publishers and content creators, librarians, open sources, students, and educators to provide copyright-intelligent links with contextual access and pricing. It is used in a wide range of universities and manages content agreements with publishers, making it easier for instructors and librarians to set up course materials and better understand how content is being used. **SIPX** also ensures that students get the benefit of all rights and permissions to which they are entitled to under the license agreement negotiated with the content producer by the university.

Lee said that it is important to make sure that any discovery system reflects user needs. Students are often curious about where the faculty get the content used in courses. Many students use a variety of search systems and supplement their use of discovery services with the library catalog. Faculty members select content using criteria such as relevance to what they are teaching, recommendations from their colleagues, cost, and ease of access. They want to find not only the content but opportunities to use it in education, and they want to be independent in choosing the content they want to use.

Analytics help in understanding how students access content, improve course designs, and increase student success rates. Use of analytics for both

discovery and ancillary services gives a more complete picture for a campus. Libraries look at subscribed and non-subscribed content that is being requested, thus helping a university to spend its budget properly. Criteria include what aspect was well engaged with by the student, how students can be assisted to understand what content is important, and at what price point will the services be used.

Publishers and Discovery Services

Nancy Blair-DeLeon, Senior Manager, **IEEE**, said that **IEEE** has been a supplier to all four of the major discovery services since 2010. All of **IEEE's** published content is available in the **IEEE Xplore Digital Library** (<http://ieeexplore.ieee.org>), and the expectation was that its usage would be maximized by discoverability because all routes would lead to it. However, the reality was that a central index was not as simple as it sounded, and a number of issues surfaced. For example, content was missing, the discovery services took content from unusual databases that did not lead back to the Xplore library, links to the full text did not work, discovery search results were different from those obtained by direct searches of Xplore, and so on.

A proactive approach was taken to resolve the issues: follow-up meetings were held; **IEEE** staff worked closely with discovery service staff; and a Discovery Services Manager was hired to work with vendors to maximize findability. Although many of the issues were quickly corrected, the overall impact of this effort is still inconclusive. For many users, Google and Google Scholar remain the services of choice, and discovery may not provide a direct route to **IEEE Xplore**. Librarians continue to lack confidence in discovery services and do not trust their relevancy ranking algorithms because they do not work well on known-item searches.

Other issues include:

- Multimedia will play a larger role in discovery in the future; what does this mean for the bibliographic record?
- Because it is difficult to track the path of a searcher from one service to another, will accurate usage reports ever be possible?
- Participation in a discovery service does not mean that it is the first and only route to the content; niche databases will continue to be important.

Despite the issues, **Blair-DeLeon** said that there are advantages and pitfalls for scholarly publishers. The main advantage is that students, especially undergraduates, become exposed to discovery services outside of their electrical and computer engineering programs earlier in their academic careers. Pitfalls include:

- More content that competes with scholarly research articles being added,
- Multimedia vs. bibliometric records,
- User interface configuration, updates, and changes, and
- Accurate usage reports.

She concluded that publishers such as **IEEE** should definitely continue to participate in discovery services.

Michael Takats, Director of Product Strategy, Web of Science, **Thomson Reuters**, said that discovery starts in many places: email alerts, RSS feeds, workflow tools, library services, and open Web search services. The Web of Science acknowledges this and has agreements with all of the major discovery services except **OCLC** to furnish them with data from journal citation indexes, giving the discovery services coverage of multidisciplinary journal data from a trusted source. Takats suggested that the most authoritative article-level metric is "times cited"; in fact, Summon uses Web of Science metadata as part of its search algorithm, and many users have high regard for records containing such data.

Customers seem to value the inclusion of Web of Science data in discovery services, but this measure is difficult to quantify. No significant increase in traffic being driven back to the Web of Science has been observed. A few minor implementation difficulties have been experienced because APIs sometimes do not work correctly, but overall, the experience has been positive, so collaborations will continue. When a user searches on "all databases," article-level data are aggregated; detailed article-level usage data are needed, which will help providers improve the user experience.

Discovering Scholarly Content

Roger Schonfeld, Program Director, **Ithaka S&R**, reported on the results of several surveys of faculty members in libraries that **Ithaka** conducted over the last two years. Faculty member practices vary; starting points of research include general search engines, library catalogs, and specific research resources. In general, researchers in the sciences are more likely to start with specific electronic resources, but humanists are more likely to start with the library catalog because they use books more heavily in their research.

continued on page 83

Library catalogs are most likely to be used for known-item searches. For current awareness, a large fraction of researchers attend conferences and use reading materials suggested by their peers. New issues of journals are also widely used. Discovery services are seen as more useful than the library in helping users find new items they had not known about.

Schonfeld listed the following points for consideration:

- How do we explain the interest in discovery services given the use of Google Scholar?
- Is there a sustainable role for the library in discovery?
- Is it possible to address all users' needs through a single search box?
- Should search results be anonymous, categorized, or personalized?

More results from the surveys are available through the Ithaka Website, <http://www.sr.ithaka.org/>.

Discovery Service Provider Experiences

Andrew Nagy, Lead Product Manager and Product Owner of ProQuest's Summon Service, noted that true discovery occurs when all content is treated equally with a unified index. Summon started with 300 million records and now has over 1.4 billion from over 100 abstracting and indexing databases. It offers synonym searching so that users can find content whether they use controlled vocabularies or not. Lessons learned during the development of Summon include:

- Meeting users' expectations is the key to success.
- Comprehensive knowledge bases must match up with collections.
- Discovery should integrate with library workflows: it is not only a tool for users but also for processes like collection development, catalog usage, and circulation.
- Discovery is about the users: their needs, how they do research, and their behavior.
- Discovery delivers value and is a success!

Besides impacting users directly, discovery has changed how libraries do collection development. Many libraries now purchase materials based on discovery usage reports, and they will not buy content if it is not covered by their discovery service. Content providers have also experienced an impact; discovery has delivered more traffic to them, especially in the area of eBooks.³

Content neutrality is critical to ensure that users can access it regardless of source. Users favor content when they can get reliable links to it, so discovery services must ensure that they do not bias one provider over another. Discovery has become more than a set of search results and is now a research tool.

Ido Peled, Director of Solutions Architecture, ExLibris, said that the common denominators of basic discovery include a single search box, faceted navigation, content coverage, citations and alerts, APIs, and extensions. Here is his list of lessons learned:

- User expectations have changed in the last 5-7 years. Initially they thought more narrowly than they do now.
- Known items must include not only a title but a full citation; a perfect title match is not required.
- Search queries have become more complex, and a single search box may not be suitable any more. Do we need to expose everything in just one search box?
- What are users expecting when they come to the library? We need to understand how different groups of users behave.
- Technology can be a game changer. Vendors who adopted new technologies have survived.
- We see more linked data now. How has it made a change in the market?

The market has evolved; we must make sure that everything the user might search for is made available. Users are now searching with different keywords and different patterns. We assume that users are searching for articles and books, etc., but many of them search for other things and are asking questions such as "who is...?", "when was...?", "where is...?" etc. The biggest market change is the emergence of unified results management, which has caused a change in library operations.

In today's mobile environment, users now expect information to be available everywhere, anywhere, and at any time. They want the ability to



turpin
distribution

A 360° SERVICE DELIVERING
GLOBAL FULFILMENT SOLUTIONS

BOOK FULFILMENT

At Turpin we understand the importance of providing a timely and efficient fulfilment service for your book customers. We support many thousands of customers throughout the global arena on a daily basis – to include key trade accounts, internet e-tailers, academic institutions and individuals. Core services include:

- Full EDI integration for key trade accounts
- Amazon Advantage programme compliant
- Bespoke P&A data feeds for improved product discoverability
- Comprehensive consolidation service for reduced shipping costs
- Dedicated key trade customer advisors
- Electronic delivery of all invoice and shipping documentation
- State-of-the-art warehouse management system providing real time access to stock levels.

SUBSCRIPTION MANAGEMENT

We can handle every aspect of the subscription fulfilment process for you, from order processing to claims handling, payment collection, renewal campaigns and label production and mailing.

We can support all types of subscription products and business models using multiple currency pricing and payment options. With over 40 years experience in the business we have developed extensive partnerships with subscription agents and online hosts to ensure there is a seamless transition from order to activation.

For more information about our services please contact Neil Castle or Bob Rooney and we will be delighted to discuss how Turpin can assist you with your print and digital distribution.

Neil Castle - Commercial Director - UK
E: Neil.Castle@turpin-distribution.com
T: +44 (0)1767 604868

Bob Rooney VP Publisher Relations - US
E: Robert.Rooney@turpin-distribution.com
T: 001 860 318 3330

personalize results on a mobile device. Recommendations have become very popular; for example, in a side panel display, Google now displays results from questions the user has not even asked. Similarly, a search interface could visualize a bookshelf and provide Amazon-like recommendations to the books on the left and right of the desired one.

Linked data is the next technology advancement that we can make. We do not think linearly and can move in any direction, so we need to be able to handle searches on broad topics. We can provide information from many different areas by indexing everything in a collection and making it available without differentiating between types of material.

John McCullough, Product Manager, Discovery, OCLC, said that WorldCat is a service governed by members and for members. It provides access to 72,000 libraries in 137 countries. From the beginning, the library catalog was about discovery. OCLC originally launched its original discovery service, WorldCat Local, in 2008, which was recently expanded into WorldCat Discovery and is now in beta test. It received 688 registrations in its first two months and now has access to 320 million bibliographic records, 2.1 billion library holdings, and 1.4 billion article citations. A new API provides access to WorldCat's central index and is available to any OCLC discovery library.

OCLC has learned the following lessons in its discovery program:

- Meet users' expectations, not just their needs, and be where they are (on mobile devices, the open Web, etc.).
- The future belongs to those who collaborate.
- Today's expectations are largely shaped by Google, which provides comprehensive results. Libraries only want to expose sources from which they can fulfil requests, which is a clash with this.
- Relevance can always be better! Known-title searching is an area for improvement.
- Provide fulfillment capability across all formats.
- The content provider's brand is meaningful to users; it should be displayed with search results. The provider's unique records and proprietary content, such as metadata, should be preserved.

continued on page 84

- The market needs deeper metadata for better relevance.
- Push the value of the open Web to users and get them to arrive at the library.

Sam Brooks, Executive Vice President, **EBSCO**, began by noting that content providers and libraries will reward vendors who are open, so **EBSCO** has made available detailed information about how relevance ranking works in the **EBSCO** Discovery Service (EDS). (Search “relevance ranking” in Google, and see the first few results.) **EBSCO** has a large group of developers working on EDS.

EBSCO has learned that the market wants collaboration from its vendors, even when they are competitors, and has therefore established partnerships and has published guidelines for additional potential partners. They have also learned that the best metadata is obtained by listening to the needs of content providers and responding to them. Other lessons have been learned from thousands of end user tests, which are continuing.

The market is becoming more savvy about how it evaluates discovery services, and a number of research studies have recently been published.⁴⁻⁶ It has also recognized the need for the best possible relevance ranking, the best content, and full OPAC compatibility, all accessible by a wide choice of front ends. **EBSCO** has responded to this desire by making its API available through nearly every major ILS vendor front end as well as through some specialized interfaces.

Content providers want to increase usage of their products while maintaining their brand identity, which means that they will not allow elements of records from several providers to be fused together into a single record. They also want access to their content to be available only to authorized users who have a current subscription to their databases. Relevance algorithms must respect the importance of subject indexing, giving priority to matches on subject headings over those obtained solely from title words or other fields. Discovery services must also display the name of the source database from which the results were obtained, along with a link to that database, allowing the user to browse the provider's thesauri and use other advanced features such as searching on words in “see,” “see also,” and similar fields. And finally, they expect a discovery service provider to refrain from misleading marketing, such as using simple title comparisons to claim coverage of their product in other indexes.

Users expect the first page of results to have the most relevant and valuable information available. **EBSCO's** view is that the best way to meet this need is to build a discovery system around high-quality subject indexes. **Brooks** said that simply extracting concepts from the full text is no substitute for employing information professionals to create subject indexes.

Besides marketing EDS to the academic market, **EBSCO** is also working with many corporations globally. Corporations have different information issues than academia, such as a high interest in patents. In general, corporate libraries, where they still exist, have more intermediaries and need more advanced functionality in discovery services.

Moving Forward: Recommended Practices

Todd Carpenter, Executive Director of the **National Information Standards Organization (NISO)**, titled his talk reporting on **NISO's** Open Discovery Initiative (ODI) “You Can't Browse the Stacks in a Digital Library” and said that standards are familiar, even though many people do not notice them. We do not wander the stacks in a digital library; instead, we have metadata-based discovery interfaces and have moved from traditional catalogs to metasearch engines and indexed search engines. We cannot separate the discovery and delivery any more — discovery without delivery leads to frustration.

According to **Carpenter**, discovery services have been implemented by thousands of libraries around the world, and they impact millions of users. Research has shown that library usage increases when discovery services are implemented. Although the open discovery environment is good for libraries and users, several issues have emerged:

- A lack of clarity about how discovery services and content providers should interact,
- Data exchange and copyright concerns,
- Concerns about ranking algorithms and bias in delivering search results to users.

NISO therefore launched the ODI for libraries, publishers, and service providers. The ODI does not get into technical issues of APIs or how

products are differentiated in the marketplace; instead, it focuses on recommendations for data format and transfer, communication of the rights of libraries in the area of indexing, definitions of fair linking, and the exchange of usage data. A survey of the discovery service community received 870 responses, mostly from libraries, and from it, a vocabulary for the systems and recommended practices for content providers and discovery service providers were developed.

Some of the recommendations of the ODI address:

- Data to be exchanged between content providers and the library community, including coverage and who should receive the data,
- Data exchanged between the discovery service and libraries, such as what fields have been indexed,
- A statement of the discovery service's business arrangements with the content providers,
- Recommended metrics to be provided to content providers and libraries, and
- The use of licenses without non-disclosure agreements, so that transparency regarding any possible bias can be maintained.

The initial ODI document was issued in November 2013, and it is expected to be published in the near future.

The future of discovery services includes the following:

- Non-textual content (audios, videos, etc.) which are increasingly being developed by researchers,
- Increasing use of datasets, software, and other new forms of science communication,
- Linked open data, particularly by library systems, and what publishers need to do,
- Altmetrics and new forms of assessment that are possible in a digital environment and how they be applied to discovery.

A white paper on the ODI was recently published by **NISO**.⁷

Carpenter concluded by noting that standards are like toothbrushes — everybody has one, but no one wants to use anyone else's. There is a variety of work in the community and we need to be aware of it and be working with it.

Jill O'Neill, Director of Professional Development, **NFAIS**, described the work of an **NFAIS** task force that developed recommended practices for discovery systems. The task force was composed of representatives from a cross section of the **NFAIS** membership. Its goal was to define the ideal recommended practices towards which the information community should strive as technology advances.⁸

There are five participants in a discovery services arrangement: the content owner, platform host, discovery service, subscriber to the service (such as a library), and users of the service. All players have multiple roles; the task force identified the rights and obligations of each role. Recommended practices for discovery services include:

- Inclusion of only contractually agreed-upon content,
- Limitation of content access to authorized subscribers and users,
- Navigation to appropriate authorized content,
- Identification of content ownership,
- Identification of platform(s), and
- Identification of content retrieved.

Attitudes towards whether discovery services are a threat or an opportunity have not shifted as much as we might think. Content providers have hesitations because they know that there are some potential threats to their existing businesses. Everyone is worried about decreasing usage, which affects library acquisitions. There is widespread concern about the loss of brand identification.

O'Neill concluded her presentation with an analogy to the 2014 World Cup matches.

In Honor of the World Cup

Technology – like a soccer ball – can be aimed to soar over our heads, land in the right spot and allow all of us to score a winning goal. But it requires teamwork and keeping our eyes on the ball – serving the needs of the user and scholarship in all its forms. 🏈

Donald T. Hawkins is an information industry freelance writer based in Pennsylvania. In addition to blogging and writing about conferences for **Against the Grain**, he blogs the **Computers in Libraries** and **Internet Librarian** conferences for **Information Today, Inc. (ITI)** and maintains the **Conference Calendar** on the **ITI Website** (<http://www.against-the-grain.com>)

continued on page 85

Back Talk from page 86

for the non-open content? How transparent can publishers be about their revenue flows, in order to explain and justify how they are pricing what they sell? Nothing like this is ever easy, and there is these days a great deal of discussion about a “total cost of ownership” approach for journal deals.

- An area of huge concern is the sensitivity of our users to matters of confidentiality and privacy. Here, license-makers need to devote a lot more attention to policies of publishers and vendors. When users enter a library system to a vendor’s site, they are leaving behind substantial amounts of personally identifiable and near-identifiable data. These vendors would not be doing their job if they did not seek to use that information to improve their product or possibly as a revenue source (sold to third parties). But library users have expectations that they will not be giving away their privacy in return for using library resources. U.S. laws lag behind those of other countries, such as European, in this regard.
- There is so much more. Issues such as perpetual licenses (what do we get to keep forever? how will that work?) and digital rights management (just how cumbersome will access need to be?) are of great importance to library users and present a landscape in constant change as technology and innovation advance. And, these days, authors’ rights are front and center — the *LMLA* attempts to support institutional authors in their right to use their own published work as they need to.

We invite you to read the *LMLA*, to think of it as a community tool, to use it as a source of support, and to improve and update it. What’s different about the corner of the library where license deals are struck is that it’s the place to see most clearly just what issues are really new, key, and/or unresolved right now, for real users, for real content providers, and thus for real librarians. By working together, we continuously improve our user environment. 🍷

Don’s Conference Notes from page 84

www.infotoday.com/calendar.asp). He recently contributed a chapter to the book *Special Libraries: A Survival Guide* (ABC-Clio, 2013) and is the Editor of *Personal Archiving*, (Information Today, 2013). He holds a Ph.D. degree from the University of California, Berkeley and has worked in the online information industry for over 40 years.

Endnotes

1. For more details, see “As Researchers Turn to Google, Libraries Navigate the Messy World of Discovery Tools,” **Marc Perry**, *The Chronicle of Higher Education*, April 21, 2014, which cites endnote #6 below.
2. <http://www.uksg.org/sites/uksg.org/files/Kortekaas%202014-06%20UKSG%20Webinar.pptx>
3. See <http://www.inforum.cz/pdf/2014/jablonka-agata.pdf> for a presentation from **Elsevier** at the 20th Annual Conference on Professional Information Resources, which reported that referrals to Science Direct eBooks from discovery services far exceeded those from other sources during the past year.
4. Comparative User Experiences of Next Generation Catalogue Interfaces, **Rice Majors**, *Library Trends*, Volume 61, Number 1, Summer 2012, pp. 186-207.
5. Paths of Discovery: Comparing the Search Effectiveness of EBSCO Discovery Service, Summon, Google Scholar, and Conventional Library Resources, **Andrew D. Asher, Lynda M. Duke, and Suzanne Wilson**, *College & Research Libraries*, Volume 74, Number 5, September 2013, pp. 464-488.
6. A Tale of Two Discoveries: Comparing the Usability of Summon and EBSCO Discovery Service, **Anita K. Foster and Jean B. MacDonald**, *Journal of Web Librarianship*, Volume 74, Issue 1, 2013, pp. 1-19.
7. Open Discovery Initiative: Promoting Transparency in Discovery, **NISO Report 19-2014**, June 25, 2014. (Available at <http://www.niso.org/workrooms/odi/>)
8. Recommended Practices: Discovery Services, **NFAIS**, August 2013, http://info.nfaiss.org/info/Recommended_Practices_Final_Aug_2013.pdf.

PRENAX
an International Company

BASCH
a Prenax Company

Subscription Management Solutions for Libraries & Corporate Procurement

- ▶ Subscription management
- ▶ E-procurement integration
- ▶ E-journal set up and activation
- ▶ E-journal URL maintenance
- ▶ Click-through access to e-content
- ▶ Cost center accounting
- ▶ Automatic claiming
- ▶ Custom and branded e-portals
- ▶ License negotiation and management
- ▶ Flexible management reporting
- ▶ Built in approval process
- ▶ Express payments to publishers
- ▶ Check in option for print titles
- ▶ Partnerships that provide usage statistics, rights management, discovery tools and single sign on.

Prenax Inc. provides subscription management solutions for procurement professionals and libraries. As a partner, we provide a single point of contact for managing electronic and paper subscriptions, professional memberships and books. We offer a true one-stop shop for all business, scientific, technical, medical, research publications and electronic content. We save you time and money and eliminate the hassle of working with multiple content suppliers.

Prenax offers the flexibility of two platforms, one for servicing libraries and one suited for serving corporate customers.

Basch Subscriptions, Inc.
Prenax, Inc.
14 Perry Street, Suite 479, Concord, NH 03301
(P) 603-225-0662 (F) 603-276-3443
www.basch.com • www.prenax.com

And They Were There from page 78

Bailey and Creibaum delivered an abridged version of their preconference program to an overcrowded room of eager listeners. The presentation highlighted the steps necessary to build an allocation formula from scratch: selecting and gathering data, determining weights, implementation, and inevitable iterative adjustments and modification. Presenters frequently reminded attendees to explore, experiment, and customize the demonstrated allocation formula to fit the unique needs of their own campus. Audience members posed questions throughout the session, many of which focused on why decisions were made to include or exclude data in the formula, possible additional factors to consider, and how the new allocation formula was received by faculty and administrators. The topic clearly lends itself to longer session format (such as a preconference program), but the presentation closely matched its program description. Attendees were directed to a demonstration formula available online for download as an Excel file and detailed presentation slides to guide their own allocation formula development. 🍷

That’s all the reports we have room for in this issue. Watch for more reports from the **2014 Charleston Conference** in upcoming issues of *Against the Grain*. Presentation material (PowerPoint slides, handouts) and taped session links from many of the 2014 sessions are available online. Visit the **Conference Website** at www.katina.info/conference. — **KS**