

June 2013

## Don's Conference Notes-Special Meeting Reports

Donald T. Hawkins  
dthawkins@verizon.net

Follow this and additional works at: <https://docs.lib.purdue.edu/atg>



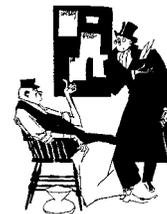
Part of the [Library and Information Science Commons](#)

---

### Recommended Citation

Hawkins, Donald T. (2013) "Don's Conference Notes-Special Meeting Reports," *Against the Grain*: Vol. 25: Iss. 3, Article 33.  
DOI: <https://doi.org/10.7771/2380-176X.6540>

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.



## In Search of Answers: Unlocking New Value From Content: The 55th NFAIS Annual Conference

NFAIS held its 55th annual conference on February 24-26, 2013 in Philadelphia, with nearly 200 information professionals in attendance. The conference program noted,

*"Each new generation of information products and services has slowly progressed along the path from delivering information towards the ultimate goal of delivering answers. Now at last the keys to actually unlocking much greater value from the content may be at hand."*

The wide variety of subjects addressed by the presenters certainly confirmed this statement.

### Keynote Address – From Databases to Networks: Knowledge in the Age of Connection

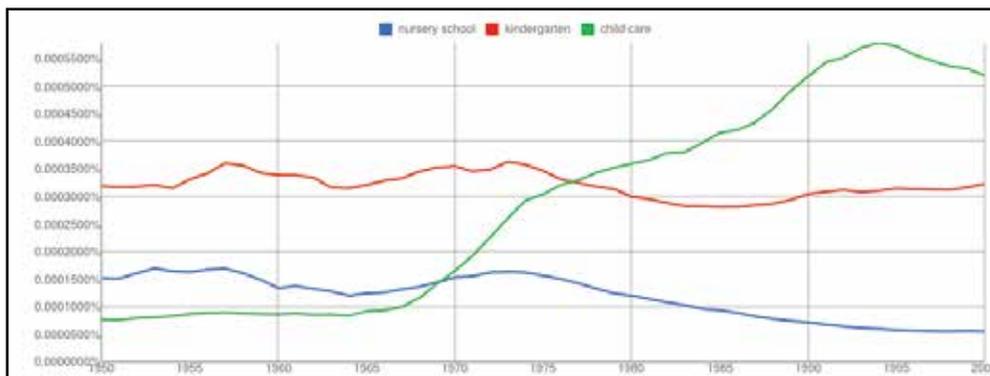
**David Weinberger**, author of the recently published book *Too Big To Know* and Sr. Researcher at **Harvard University's** Berkman Center for Internet & Society, began his keynote address by noting that our brains are small and the world is very big, so we have adapted a strategy of reducing the amount of information we deal with, a move that libraries have helped by developing advanced information filtering tools. But now we have a new medium — the network — and when we get new data, we can find it immediately when it arrives on the Web. A vast network of data, documents, comments, etc. has emerged. But does all this information have value? **Weinberger** described three ways to squeeze out more value from information.

1. *Iteration.* Software developers have created the best learning environment in history, and the chances that you are the first person to encounter a problem are vanishing. So you can pose a question and be virtually assured of getting an answer. The result is a tremendously productive environment.
2. *Platforms.* Value is generated at the top layer, and work is done below. Content is at the bottom. Today there is no difference between data and metadata; the only difference is in functionality and how it is used. On a platform, no additional value is wasted.
3. *Commons.* Knowledge and data are getting linked. The consequence of a linked data commons is that research can find vast amounts of data of many different types, and we can now easily get what we need. Linked openness enables value to scale.

### In Search of Answers: Scholars and Scientists Speak Out

The ultimate goal of online search systems has long been to provide answers, not just lists of references, to their users. We are now in an environment of severe information overload. Many scientists and researchers are under significant time pressure and are finding searching frustrating and difficult. This session featured three speakers from academic institutions describing the problem and suggesting solutions.

**Dr. Timothy Hitchcock** from the **University of Hertfordshire** in England said that the humanities are marked by ambiguity, making it difficult to move from data to knowledge and resulting in "drive-by scholarship": simply finding a quote to form the basis for a peer-reviewed article, which leads to "intellectual junk food." But the humanities are now changing rapidly as they move towards big data and textual analysis. **Google's Ngram** viewer (see <http://books.google.com/ngrams/info>) has been a game changer. It allows one to trace the occurrence of terms over time, which has spawned a new discipline: "culturomics." The figure below is an example of an **Ngram**; it shows the frequency of the terms "nursery school," "kindergarten," and "child care" in books published in the U.S. between 1950 and 2000.



New tools are allowing innovative uses of the **Ngram** viewer, such as for correlation analysis. Mapping programs such as **Google Earth** have made it easy and free to explore spatial dimensions. Digital text reading and analysis tools like **Voyant** (<http://voyant-tools.org/>) provide the ability to conceptualize text as it is being read.

The world's largest electronic flow of information is speech, not email as many people might imagine. **John Coleman**, Director of the Phonetics Laboratory at **Oxford University**, described **AudioBNC** (<http://www.phon.ox.ac.uk/AudioBNC>), a snapshot of spoken British English compiled in the early 1990s

that contains about 4,000 text samples with a total of about 100 million words. Problems working with databases of audio material include:

- Finding data is hard. How does a researcher find segments of interest? How can the data be marked up to facilitate browsing? How can large-scale audio collections be made accessible?
- Getting the data. Just copying a year of data takes a day. We need tools for browsing, searching, saving, and linking to clips of words.
- Sharing. Standards need to be agreed on and followed by compilers and users.

### Unlocking New Value from Content

This session examined how analytics are being applied by content providers in order to find new value from their content. **Ronald Snyder**, Director of Advanced Technology at **JSTOR**, described a project to mine **JSTOR's** data. Even though it is heavily used (10 billion page views, 1.25 billion searches, and 580 million PDF downloads to date), **JSTOR** has an incomplete and a fragmented understanding of its users. User expectations and a more competitive environment have increased the need for analytics. A data warehouse project initiated in 2010 has resulted in new personalized user-centric features and improved access for users.

**Stephen Abram**, formerly Vice President at **Cengage Learning**, gave another of his provocative and thought-provoking presentations, challenging users to "Get off the Hit Train." He said that we have had a bad addiction to counting clicks, which do not measure anything that matters to an end user of our data. We need to start talking to our customers about whether learning happened.

What are we counting and sharing? Titles, clicks, downloads, sessions, and session lengths are some metrics commonly used today, but do they tell us whether the user found the right article or whether learning happened?

*continued on page 69*

Was there an impact on research or strategic directions? Did we impact discovery, creativity, patents, etc.?

Here are some fascinating statistics that should be used to influence the development of new products.

- 27% of our users are under 18; 59% are female, and 29% are college students.
- Only 5% are professors and 6% are teachers.
- Every day, 35% of our users access our databases for the first time, and only 29% of them found the databases via a library's Website.
- 59% found what they were looking for on their first search.
- 72% trusted our content more than Google, but 81% still use Google!

**Abram** challenged the audience whether they could imagine a worse experience than simply being dumped into a collection of millions of articles. Can we create a good user experience? Statistics do not tell us what actually happened. We must align our databases with users' learning styles, which are heavily visual, and until we do this, we will not be building products correctly.

**Chris Burghardt**, VP, Product and Market Strategy, **Thomson Reuters IP and Science**, described a study which found that many researchers deposit their data in institutional repositories or make it available on a personal Website. There are no universally accepted standards for citing data, and many researchers feel they are not being receiving appropriate credit for their work. The result is inefficient and inconsistent citation behavior.

**Burghardt** said that the digitization of data has created tremendous opportunities to share all types of research data. We can start to take advantage of data repositories by enabling the discovery of them in the context of traditional literature and by establishing standards for attribution as well as incentives to make data discoverable. Funding mandates recently issued by the **National Science Foundation** and **National Institutes of Health** require grant recipients to deposit their data in a repository, which has spurred the establishing of over 500 data repositories. They are tracked at <http://databib.org/>.

**Thomson** has created the **Data Citation Index**, a new database on the **Web of Knowledge** platform that provides a single point of access to quality research data from repositories in a variety of disciplines. Criteria for including a repository in this database include subject relevance to the content, stability of the repository, and links from the data to the research literature. The **Data Citation Index** is expected to aid in the discovery of data most relevant to scholarly research and link it to the published literature.

**Helen Parr Moran**, VP, Smart Content Strategy, **Elsevier Health Sciences**, said that

in healthcare, one of the major challenges is getting the right information to doctors so they can make the best decisions and provide the best clinical care. **Elsevier's** approach is to use semantic enrichment in the delivery of smart content. The **Elsevier Merged Medical Taxonomy (EMMeT)** includes acronyms and abbreviations to help make search more intuitive. Its **ClinicalKey** system (<https://www.clinicalkey.com/>) includes an ontology that helps to focus results by retrieving the most relevant paragraphs from articles.

### Lunch Address – The End of the Middleman. Predicting the Future of Education Publishing

**Michael Cairns**, President, **Information Media Partners**, noted that content platforms are the future of education publishing. Platforms are nothing new in the publishing industry, but some publishers have difficulty with the idea that their platforms must be extensible.

Huge changes have been caused in the industry by **Amazon** and the Kindle. **Cairns** predicted that more big changes are coming in the trade business, and perhaps another major merger (**Cengage** and **McGraw-Hill**?). Current trends supporting his view are a slowing of growth in eBook sales, e-readers being replaced by tablets, students' dislike of e-readers for textbooks, and elimination of competition caused by mergers. Will the current business model collapse as a result? Most old-time industries are tradition-bound, and education is no different. The value chain is compacting in all segments of the publishing industry and is changing the financial model for many publishers. Platforms will come into the education space; **Pearson** has been working hard at this for at least six years and is likely to become more aggressive.

Four elements integral to a platform approach are:

- A branded content model,
- Workflow solutions and technologies,
- A common taxonomy and ontology, and
- A consistent revenue model.

Higher education is straining to prove its relevance; a recent survey showed that 45% of students had no significant knowledge gains after two years of college. The manner in which education is delivered and content is produced is under stress. Higher education has been rocked by the concept of the Massive Open Online Course (MOOC), in which there are no physical limitations of a campus. The custom textbook market is growing faster than the overall market. New companies are developing degree programs correlated to business, and MOOCs present major opportunities for this, especially in markets like China and India.

In a provocative article in *The American Interest* (<http://www.the-american-interest.com/>



*David Weinberger, the opening keynote speaker. All photos courtesy of Donald T. Hawkins.*

[article.cfm?piece=1352](http://www.the-american-interest.com/article.cfm?piece=1352)), **Nathan Hardin** predicted that in 50 years, half of the 4,500 U.S. colleges and universities will be out of business. Opportunities for innovators will emerge, particularly in the merger of the education and business markets. **Cairns** concluded by noting that we are seeing a compacting of the value chain and the ability of content producers to reach users directly, which would have been impossible in the old world where a middleman is necessary. There is an exciting future to come in educational publishing, and we are only at the start of it.

### Unlocking Content Value: New and Emerging Technologies

**Daniel Mayer**, VP of Marketing, **Temis, Inc.**, echoed a theme that was common at this conference: content users are crying out for help because digital content is doubling every two years. One of the ways to help them is by semantic content enrichment, which includes structuring of information by deriving metadata and facets to narrow query results, and the insertion of links to related information to make it easier to find similar related documents. Semantic content enrichment has a 3-fold value proposition:

1. Compelling content,
2. Semantic products, and
3. Editorial productivity.

**Alberto Pepe** from the **Center for Astrophysics at Harvard University** used the **Astrophysics Data System** as a case study to illustrate the use of new technologies to unlock content value. We are producing 21st-century research results that are written using 20th-century tools (i.e., **Microsoft Word**), and packaged in a 17th-century format (paper). This way of publishing is not working any more. **Pepe** has built a version control system, **Authorea** (<https://www.authorea.com/>), from open source technology that enables writing articles directly in the browser and sharing them interactively. (Most people now collaborate by sending **Word** files back and forth.)

### Measuring Value with Metrics and Analytics

This session explored how the value that is being unlocked from the content providers' products can be measured. **Andrea Michalek**, Co-Founder of **Plum Analytics**, asked "Are Alternative Metrics Still Alternative?" Not all uses of information generate citations. Well known problems with citation analysis include self-citations, citations added because of pressure from editors, and negative citations. Different measures of impact coming into prominence include usage, captures, mentions, and social media in addition to citations. Recency plus impact is a critical metric.

*continued on page 70*

The new challenge is in knowing all places where an article could be published — Websites, traditional journals, open access journals, institutional repositories — and merging metrics from all those sources. How much does the world use your research? Competition for research dollars is at an all-time high, and governments are demanding efficiencies in the programs they pay for through grants. Benefits of using these new metrics for content providers are: identifying new collections and fields of study, uncovering new authors in a field, the ability to target marketing efforts and measure their success, enhancement of their products, and increasing ad revenues. The altmetrics are out there; we need to find and use them.

**Caitlin Trasande**, Head, Research Policy, **Macmillan Publishers**, said that research organization administrators can use alternative metrics to help them measure performance. Examples are:

- Tracking and indexing publications. **Nature Publishing** expanded the index of *Nature* by establishing an Asia-Pacific edition and promoted awareness, access, interpretation, excitement, and transmission.
- Tracking media citations. **Altmetric** (<http://altmetric.com>) counts links to DOI-containing articles and provides tools to track the online attention paid to scholarly articles.
- Taking credit for all of an institution's research. **Figshare** (<http://figshare.com>) allows researchers to publish all their results (including videos and datasets that are usually placed in a supplemental data section) in a shareable form quickly on any desired platform.

### Closing Keynote – In Search of Answers: IBM's Watson and a Look at 2020

How can a computer understand the meaning of strange questions? **Frank Stein**, Director of **IBM's Analytics Solution Center**, described some of the lessons learned in the development of **Watson**, the famous computer that was able to beat a chess master and win at the

Jeopardy quiz program. We are drowning in information and need the tools to turn 4 zettabytes (4 billion gigabytes) of information (4 times as much as in 2010) into insights.

The goal that **Watson's** designers set themselves was to design a computing system that rivals a human's ability to answer questions posed in natural language by analyzing vast amounts



*Robert Snyder as he receives the Miles Conrad Award from NFAIS President Barbara Dobbs Mackenzie.*

of information in real time. In contrast to chess, real language is very difficult to analyze because there are a seemingly infinite number of ways to express something, and language is ambiguous, contextual, and implicit. The challenge faced by **Watson** was not only understanding questions but finding and validating answers with confidence and high precision, and doing it fast enough to compete on Jeopardy. It is hard to program a computer to instantly know with confidence what is in the answer set and what is not.

The Jeopardy data covers a huge range of topics, so the developers had to gather data from a multitude of sources such as every version of the *Bible*, the entire Wikipedia, complete works of **Shakespeare**, etc. Sub-questions were constructed to answer the main question, and a model to give the best possible answer based on 10,000 previous Jeopardy questions was built. The model used semantic networks, lexical calculations, and syntactic relationships to find hidden associations between seemingly unrelated terms. The Jeopardy system proved a point, but not many companies want to buy a computer to play Jeopardy!

**Watson's** developers are now moving to a multi-dimensional processing system to solve problems in healthcare, which has some of today's most complex information challenges. The amount of medical information is doubling every five years, and most doctors have little time to grasp all the new literature. Two applications are being developed for **Watson**: utilization management (ensuring that what is prescribed is based on evidence), and oncology diagnosis and treatment.

Because of space limitations, descriptions of several of the presentations could not be included in this article. Slides from many of the presentations are available on the conference Website at <http://nfais.org/event?eventID=399>.

*continued on page 71*

### The Miles Conrad Memorial Lecture

A highlight of the annual **NFAIS** meeting, the **Miles Conrad Award** is the highest honor bestowed by the **Federation**. Established in honor of **NFAIS** founder and first president **G. Miles Conrad**, the award is given to an innovator and leader in the information industry. The list of previous award recipients reads like a "Who's Who" of the industry. The award consists of a cash honorarium and plaque. This year's recipient and lecturer was **Robert Snyder**, co-founder of the **Cambridge Information Group (CIG)** in 1971 and current Chairman of the Board. **Bob** was owner and Chairman of **Cambridge Scientific Abstracts (CSA)**, which was merged into **ProQuest, LLC** when **CIG** purchased **ProQuest**, and of **Disclosure, Inc.**, which is now owned by **Thomson Reuters**. He is currently a member of the **Smithsonian Institution** Libraries Advisory Board and several other boards in the Washington, DC area.

**Snyder's Miles Conrad Lecture** followed the conference theme and was on the subject of unlocking the future of knowledge businesses. He said that what users want and what they will always want has not changed. **CSA** was created to access scientific information and to enhance knowledge. It was not defined as a creator of abstract journals, but emphasized the delivery of tools to add value to the research process.

**Snyder** traced the history of **CSA's** use of technology as it evolved, from manual typewriters, then **Wang** word processors, CD-ROMs, and finally, the Internet. Each step led to efficiencies in production. He described how **CSA** has pioneered fast, comprehensive searching, which with the decline in search costs has opened up inexhaustible possibilities for accessing information, and it has not wavered from that commitment. One of its newest services, the **Summon** discovery service, reveals the breadth of libraries' collections and is still evolving.

**CSA** followed the work of **Dialog** and was one of its early database suppliers. This was online searching before the days of the Internet! **Dialog** is now **ProQuest Dialog** and is still a precision tool for complex searching. It in the process of being repurposed and is making an impact on the industry again.

In **Snyder's** opinion, eBooks will be ever more important to searchers; thus, **ProQuest** has acquired **E-Book Library (EBL)** and will merge it with the previously acquired **ebrary** to provide a single eBook platform for all of **ProQuest's** services. **ProQuest** has acquired many companies over the years, and its vision is to make everything work together so that silos are gone and search becomes more comprehensive. In 2013, **Summon** and **ProQuest** will be fully integrated.

**Snyder's** passion is faster, cheaper, and more comprehensive search, and we need to think about improvements to it by responding to the needs of the researcher. We will not survive by building fiefdoms to search only our own content; instead, we must connect researchers to the best and most relevant information so that a single query can tap all resources at once and introduce the researcher to sources previously unknown to them.

## Information Discovery and the Future Role of Abstracting and Indexing Services: An NFAIS Workshop

A common misperception in today's era of widely available information of every type is that there is no longer any need for abstracting and indexing (A&I) services. This **NFAIS Workshop** in Philadelphia on March 15, sequel to a previous workshop on information discovery (see *ATG v.24#6, December 2012 - January 2013*, p.66), made it very clear that A&I services still have a valuable role to play. It began with a review of history and then explored these services' role from a librarian's view, looked at two emerging players, examined how existing A&I service providers are adapting to the current information environment, and concluded with some predictions for the future.

### History, Mission, and Current Status of A&I Services

**Bonnie Lawlor**, Executive Director of NFAIS, traced the role of A&I services throughout history. We are acutely aware that information overload is a significant issue today. But it is not new. Overload began as soon as printing presses started producing publications; in fact, the first journal (*Journal des Scavans*) published in 1665 contained abstracts of articles. Formal abstract journals began in 1820 with the *Pharmacopeia of the United States*, and their number has grown exponentially since then. According to an article published in *Learned Publishing*,<sup>1</sup> in 2009, an estimated 1.5 million articles were published that year. A&I services are an answer to this explosion of information.

During the 1950s, a huge increase in the number of journals significantly impacted A&I services. *Chemical Abstracts* fell three years behind, for example. The only way to cope was for the services to embrace computer technology, and they became very early adopters of it, which gave them an initial competitive advantage in the digital age. Indirect benefits of computerization included production efficiencies, increased currency of content, and opportunities to develop new products. As a result of the move to electronic processing, business models changed: print sales declined, and licensing of databases became more common. Electronic products are intangible, so product loyalty was hard to maintain. Staff had to be retrained in new skills such as customer service, and markets became global. And digital content requires an ongoing investment in system upgrades.

Price increases have resulted in the emergence of alternative publishing models, leading to today's emphasis on open access publishing and the mindset that information should be freely available, which has been encouraged by funding agencies' mandates that research results must be deposited in freely accessible repositories.

Searching behavior has dramatically changed, and now a search engine is the first choice of many users who rely on themselves to select results. Many of today's students there-

fore equate research with using **Google**. Users now want convenience, linking, interactive search systems, easily discoverable supplemental material, and analytic tools, all packaged in a pleasurable and reasonably-priced (or free) search experience. Mobile phones and social media have had similar far-reaching impacts on A&I services; delivery of information to hand-held devices has now become the norm.

In the face of these major changes, will A&I services survive? As long ago as 2003, **F.W. Lancaster** asked this question and concluded, "the viability of a vast network as an information resource must depend upon the imposition of quality filters similar to those in a print-on-paper world."<sup>2</sup> These filters will be A&I services and will continue to have a strong role. The ultimate decision lies with the user; whatever serves their needs will go forward.

### Role of A&I Services in Information Discovery: The Librarian's Perspective

**Lawlor's** opening presentation was followed by a session in which three librarians discussed how their users search for information and the role that A&I services play. (Librarians are major users of A&I services.) **Andrew Asher** from **Bucknell University** conducted a study of 86 students who used a discovery service. Students have become used to **Google's** single search box, and they generally do only simple searches with one or two keywords. If they do not find what they are looking for after a cursory evaluation of the search results, they tend to start another search with different terms rather than refining the one they began with. Most students do not understand how searching works and tend to use poorer quality terms, assuming that if something was not found, it does not exist.

Students become very loyal to a tool that works for them, even if it does not contain the most appropriate databases for their search. They like discovery services and trust them to retrieve the best results. They believe that the first five to ten results are the best (or only) ones the library has to offer. Developers and librarians must therefore be careful in setting up the defaults for discovery tools because of their effect on search results. Detailed results of this survey will be published in a forthcoming article in the July 2013 issue of *College & Research Libraries*.

According to **Chris Strauber** from **Tufts University**, full text is ultimately the point of the search process, but how one gets there is also important. A&I services can be superior resources because their indexes are compiled by people who understand the subject area, and the metadata is in the language of the discipline. Discovery services are good for exploratory questions, but they have limited metadata. Indexes and abstracts add human expertise to what computers can do. The

browsing function is just as important as search, particularly for the humanities; lots of questions can be answered simply by browsing.

The final presentation in this session was a report on a white paper produced for **Sage Publications** and presented at the **ALA Midwinter** conference in 2012.<sup>3</sup> It discussed best practices for access and discovery of content in libraries, as well as problems that libraries, publishers, and vendors need to solve. Cross-sector collaboration is necessary in the discovery of scholarly content, and collaborative groups such as the **NISO Open Discovery Initiative (ODI)** are being formed to develop standards and best practices for pre-indexed library discovery services. Linked open data has become very prominent, and the open metadata concept is growing in popularity.

Librarians and publishers can add value for learning by integrate their expertise into user workflows. For example, **Purdue University** has appointed a "data services librarian" to help with grant writing and meeting funders' requirements. But libraries and publishers have not provided a unified user experience because of different fulfillment options, metadata models, etc. Legacy databases must satisfy users' needs for content on a variety of devices. A&I services cannot rest on their laurels and continue to depend on growth in their markets. They must invent new ways of explaining their value proposition and of participating in the semantic Web.

### Emerging Players in Information Discovery

Representatives of two new players in the discovery services area, **Molecular Connections** and **Mendeley**, described their products. **Molecular Connections** aggregates content from Web sources into a coherent database. It is the largest A&I company in India and operates in three major areas: mining, representation, and creation of content, particularly in biological and pharmaceutical areas. Its product, **MC-Outlink** (see <http://www.molecularconnections.com/publishing/en/home/publishing-offerings/mc-outlink>) obtains relevant information that is dispersed across Web sources, current news, videos, etc., in addition to published scientific content and creates a report in a standard format. **Jignesh Bhate**, CEO, estimates that to gather all the relevant information related to a single drug would require two to three hours or more; thus, **MC-Outlink** can be a major time saver for researchers.

**Jan Reichelt**, Co-Founder and President of **Mendeley**, described how his service extracts data and full-text information from a wide variety of sources, annotates it, and aggregates the information in the cloud, thus creating a social layer between people and their research interests. The most relevant content is then pushed to the researcher. For groups of researchers, relevant articles can be sent to members of the group who can then discuss them in a manner similar to **Facebook's** news feeds. Information is anonymously aggregated and combined into users' social environments without breaking their privacy.

*continued on page 72*

What would happen if there were no boundaries around social sharing, so that information from the Internet was made visible to others? Additional revenue sources might be created for information providers. Value is created through sharing and embedding, enabling the community to connect. Some services, such as **kleenk.com** and **openSNP**, have begun developing products using **Mendeley's** data.

### The A&I Services' Perspective of the New Information Landscape

In a presentation entitled "A&I Services: Enhanced Relevance through Aggregation and Discovery," **Craig Emerson**, VP, Publishing, **ProQuest**, said that much of the A&I business model has not changed, but it goes through times of disruption. Indexing and tagging must be very good to deal with these times, which are caused by a rapid increase in the number of publications, open access content, personal repositories, and article-by-article publishing. The A&I services must also compete in the face of published articles saying that **Google Scholar** is good enough for literature reviews. Comprehensiveness is important; a recent article noted that the cost of missing an article could be up to 76 days and \$10,000.

Because A&I services are the starting point for much academic research and remain the first choice of many researchers, they are holding their own and are still being heavily used. Editorial relevancy adds significant value. The changes in content such as new fields (companies, people pictures, materials, document types, etc.) and the need for deep indexing of figures, tables, and datasets are challenging but necessary to add search precision. Summarization is attractive to many users do not have time to read complete articles and simply scan through them. Video curation is also becoming more significant.

Libraries are widely recognized as a superior source of quality content, but there is a general lack of awareness of such resources. Discovery services turn the complexity of a library site with lists of databases into a **Google-type** approach with a single search box. Discovery services and A&I services are serving two separate needs: A&I services provide precision discipline-specific searching for expert researchers, and discovery services provide quick access to full text. Both approaches are necessary; however, convenience will always trump content. **ProQuest's** Summon service has increased overall use of library resources significantly, and 60% of student users said that it "improved their ability to do research." It also had a major impact on usage of A&I databases.

**Lynn Willis**, Content Development Manager, **American Psychological Association (APA)**, concurred with **Emerson** and described how similar steps have been taken to improve **APA's PsycINFO** database. New fields, such as Access URLs, DOIs, author email addresses, and cited references, were added. To cope with the explosion of content, machine-aided indexing technology has reduced indexing

time by about 50%, and new databases for grey literature, psychological tests, questionnaires, and computer programs were developed.

**Roger Schenk**, Content Planning Manager, **Chemical Abstracts (CA)**, noted that **CA** is now almost exclusively delivered electronically and is rapidly moving into mobile delivery. It currently covers 63 patent authorities and over 10,000 journals. Much of the current literature growth comes from patents, and China is a major driver. Challenges include currency, timeliness, budgetary stresses of customers, and competition from free government databases and search engines like **Google**. A&I services must balance their challenges with those of their customers.

Content innovation and technology have significantly simplified scientific literature searching and have provided a new area of opportunity: evaluation. So **CA** has added more context and relevance to every patent and journal article abstract. Recently, graphical abstracts (chemical structures and illustrations) were added, as well as links to the full text.

**Ryan Bernier**, Director, Subject Indexes, **EBSCO Publishing**, emphasized that **EBSCO's** business is subject indexes, and their full text products mostly began as A&I databases. Subject indexes are a necessity, and the quality of searches depends directly on the quality of the indexes.

**EBSCO** is looking at several ways to index and abstract non-textual content because the use of additional materials such as datasets and images is growing. The **Associated Press** is working with **EBSCO** to add images, graphics, and sound bites to the **EBSCO Discovery Service (EDS)**. Open access journals are also being aggressively added, and more records are linking directly to freely available full text.

**EBSCO** does not contribute to any discovery service, which is standard practice for many subject index providers. It has its own discovery service, but that is only a small part of **EBSCO's** business. Customers of both **EBSCO's** search service and **EDS** can opt for platform blending, making both services work seamlessly together. The main focus is to ensure that subject indexes thrive.

### Information Discovery and Future Players

**Carl Grant**, Associate Dean, Knowledge Services, **University of Oklahoma**, concluded the workshop with an excellent presentation on the future of information discovery and the roles that A&I services might play in the future. He listed some of the top trends in the information industry today:

- Mobile apps will increase.
- More and more data will be stored in the cloud.
- Private enterprise app stores will appear and will exert more control over the data. How will we interface with them?
- The Internet of Things will emerge, with sensors in mobile phones. Not many apps are available yet in this area.

- Single warehouses of big data will be abandoned in favor of silos of big data.
- In-memory computing will enable real-time analysis and transactions.

Discovery interfaces have a large market share; only 11 **Association of Research Libraries (ARL)** institutions are not using them. Libraries have lost a lot of ground, and as the number of librarians decreases, the need for discovery will grow. Delivery has become the core business of libraries, not discovery. Our territory is being lost even as we think we are defending it. An excellent analysis of discovery in the world of libraries by **Lorcan Dempsey** appeared in **EDUCAUSE Review** and should be read by all librarians.<sup>4</sup>

Usage of mobile devices as access devices has now passed that of PCs; in fact, one author predicts that many of today's young people will never own a PC because a tablet will be all they need. Is all of your content available in the cloud? Do not underestimate the importance of the unbundling of education and the appearance of MOOCs (massive open online courses) as another point where your data must be available. Content must be deliverable through a variety of platforms — HTML5, APIs, and Web services. Users are now constantly connected to the Web; messages have become shorter; and so have attention spans. Learning styles are 29% visual, 34% auditory, and 37% tactile, which has implications for content delivery. Unfortunately, we tend to ignore all but the visual.

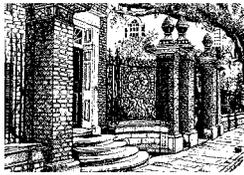
Change will continue, so what is the role of A&I services? We must think bigger! Here are some ideas:

- Create learning courses out of abstracts so people can take a quick course from them.
- Think about multiple languages.
- For your content to be discoverable, you must provide support for APIs and Web services so it can be widely integrated in numerous delivery platforms.
- Index and abstract far more than just printed works.
- Be sure to index open access journals — there is a huge move toward them.
- Realize the amount of content created by individuals — it will account for nearly 70% of the digital universe in the near future according to **IDC**.
- Build user profiles so you can deliver services based on their needs.
- Video is growing rapidly.
- Add datasets into current practices.
- Give students a pathway to deep content by indexing deeper into the Web.
- Think mobile. Base delivery of information on sensors in mobile phones. Don't try to dumb down the interface and just squeeze everything onto a smaller screen.

*continued on page 73*

Subscribe to The Charleston ADVISOR Today!

6180 East Warren Avenue • Denver, CO 80222  
Phone: 303-282-9706 • Fax: 303-282-9743



# The Charleston ADVISOR

Critical Reviews of Web Products for Information Professionals

*"The Charleston Advisor serves up timely editorials and columns, standalone and comparative reviews, and press releases, among other features. Produced by folks with impeccable library and publishing credentials ...[t]his is a title you should consider..."*

— *Magazines for Libraries, eleventh edition, edited by Cheryl LaGuardia with consulting editors Bill Katz and Linda Sternberg Katz (Bowker, 2002).*

- Over 750 reviews now available
- Web edition and database provided with all subscriptions
- Unlimited IP filtered or name/password access
- Full backfile included
- Comparative reviews of aggregators featured
- Leading opinions in every issue

**\$295.00 for libraries**  
**\$495.00 for all others**

**Yes! Enter My Subscription For One Year.**     Yes, I am Interested in being a Reviewer.

Name \_\_\_\_\_ Title \_\_\_\_\_  
 Organization \_\_\_\_\_  
 Address \_\_\_\_\_  
 City/State/Zip \_\_\_\_\_  
 Phone \_\_\_\_\_ Fax \_\_\_\_\_  
 Email \_\_\_\_\_ Signature \_\_\_\_\_

## Don's Conference Notes from page 72

- If you are not directly facing the user, make sure that your APIs can do that. Indexing is the key for filtering.
- Find your unique value contribution — the days of the average are over.<sup>5</sup>
- How can all these enhancements be done without hiring all the necessary staff? Some libraries have enlisted users to help in content creation: creating tags, for example.
- China is coming at us like a freight train! They are starting to build digital libraries from the ground up. 🌸

*Donald T. Hawkins is an information industry freelance writer based in Pennsylvania. 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.infotoday.com/calendar.asp>). 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. **Jinha, Arif E.**, *Learned Publishing* 23(3): 258-263 (July 2010) (available at <http://alps.publishing.ingentaconnect.com/content/alps/lp/2010/00000023/00000003/art00008>).
2. **Lancaster, F.W.**, "Does Indexing and Abstracting Have a Future?" *Anales de Documentation*, No 6, 137-144 (2003).
3. "Improving the Discoverability of Scholarly Content in the Twenty-First Century," (available at <http://www.sagepub.com/repository/binaries/librarian/discoverability-whitepaper>).
4. **Dempsey, Lorcan**, "Thirteen Ways of Looking at Libraries, Discovery, and the Catalog: Scale, Workflow, Attention," *EDUCAUSE Review Online*, January/February 2013, (available at <http://www.educause.edu/ero/article/thirteen-ways-looking-libraries-discovery-and-catalog-scale-workflow-attention>).
5. **Friedman, Thomas L.**, "Average is Over," *The New York Times*, Editorial, January 25, 2012, (available at [http://www.nytimes.com/2012/01/25/opinion/friedman-average-is-over.html?\\_r=0](http://www.nytimes.com/2012/01/25/opinion/friedman-average-is-over.html?_r=0)), and "Average is Over, Part II," August 7, 2012, (available at <http://www.nytimes.com/2012/08/08/opinion/friedman-average-is-over-part-ii.html>).

## Biz of Acq from page 67

their input in developing usage reports; this includes involving librarians in usability testing. Similarly, librarians must be willing to engage in this process not only with vendors, but also in making improvements to COUNTER and other standards. 🌸

## ATG Crossword Puzzle from page 14

