

2014

# Op Ed: Embracing the Digital: Libraries and the Evolving eBook Ecosystem

Mark Dahl

*Aubrey R. Watzek Library, Lewis & Clark, dahl@lclark.edu*

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



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

---

### Recommended Citation

Dahl, Mark (2014) "Op Ed: Embracing the Digital: Libraries and the Evolving eBook Ecosystem," *Against the Grain*: Vol. 26: Iss. 4, Article 11.

DOI: <https://doi.org/10.7771/2380-176X.6798>

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.

# Op Ed — Embracing the Digital: Libraries and the Evolving eBook Ecosystem

by **Mark Dahl** (Director, Aubrey R. Watzek Library, Lewis & Clark, Portland, OR 97219) <dahl@lclark.edu>



The **Oberlin Group's** recent call for library rights to interlibrary loan eBooks calls attention to the challenge of sharing eBook content between academic libraries. Indeed, academic libraries must figure out a way to continue to deliver on patron expectations for broad access to book content in an evolving eBook ecosystem.

The successful strategy for doing this will harness the unique advantages of the digital environment through experimentation and negotiation between libraries and publishers. Though the **Oberlin Group** statement calls for libraries and presses to “work together,” the “principles” that it presents at the end of its statement do not suggest a flexible approach but rather a rigid one that attempts to apply the old rules for sharing in the print environment to the new digital one.

Though mandating that “there should be no limits and no additional costs related to the number of times an eBook can be accessed over time” might seem ideal from a library and a patron perspective, such a requirement would limit the purchasing models that publishers and book vendors could offer libraries. It could very well lead to a rigid and exclusionary marketplace with higher prices for eBooks and fewer options for libraries.

Academic library patrons have come to appreciate the access to a wide swath of book content that modern interlibrary loan systems provide them. In the print world the copyright first-sale doctrine lets a physical copy of a book be used by an unlimited number of parties. The internet and modern discovery-to-delivery systems have allowed libraries to maximize the first-sale doctrine by moving physical books quickly to readers that need them.

In the digital world of eBooks, libraries (and individuals) lose the first-sale doctrine, but also pick up advantages that don't exist in the physical world: the possibility of leasing access to large virtual libraries of content, the lack of physical space required for eBooks, the elimination of costs associated with managing and tracking materials, and the possibility of just-in-time purchasing.

The emerging eBook ecosystem is not perfect, but with library and publisher negotiation and cooperation it is evolving and getting better. Going forward, academic libraries are likely to offer a wide range of digital books to their patrons through a variety of purchasing models including licensed eBook packages, on-demand purchasing, open-access eBooks, and in some cases, interlibrary loan.

## Licensing Collections

As media consumers in the year 2014, we're familiar with a marketplace that continues to evolve in response to changes in digital media. Think about how we now access music, films, or television. We used to get movies at the video store, but then **Netflix** broadened our horizons with a vast library of DVDs available for delivery, including a “long tail” of less popular but potentially high quality content. Now subscriptions to **Netflix** or **Spotify** provide access to massive digital video and audio libraries instantly, and we supplement them with individually purchased content.

Leased eBook packages such as **ebrary's** Academic Complete are akin to a Netflix streaming subscription. They provide academic libraries access to a large collection of content for a low recurring cost. They are a good way for libraries to get started in the eBook arena at an introductory price, especially when compared to building a print collection of similar size.

Licensing these collections provides no provision for long-term access to the titles and no rights to interlibrary loan, but the trade-off is reflected in the low cost. Many librarians criticize them because they lack the ownership rights that are characteristic of print collections. But in fact they represent the digital marketplace providing products that offer substantial value to libraries.

## Pay-as-you-go

Recently published eBooks are usually not included in these discount packages and typically must be purchased separately at retail price or via more expensive package deals. But the digital environment has given libraries a big lever when it comes to acquiring this type of eBook: the possibility of purchase on demand.

Libraries have traditionally expended copious amounts of space, human resources, and acquisitions dollars on print journals and books that wait in the stacks just in case they are needed (the likelihood of a book being used once at a research library has been found to be 50-60%). In the digital world it's possible to provide library patrons the ability to search through an extensive catalog of books and only pay for what they actually access.

Demand-Driven Acquisitions for eBooks is an evolving practice where a library provides patrons with a wide cross section of books to search. Without the patron's knowledge, when they access a book their action triggers short term loans and eventually a purchase by their home

library. Libraries are still negotiating with eBook vendors about the actual terms of what triggers a purchase.

The **Orbis Cascade Alliance Consortium** in the Northwest is slowly building a shared collection of eBooks using such demand-driven purchasing mechanisms. Shared ownership of the materials avoids the sticky question of interlibrary loan, though it's a challenge to scale the project due to the wide variety of acquisitions goals and budgets within the consortium.

Some libraries have been exploring ways to reshape the book publishing landscape by contributing to open access eBook initiatives. The open access approach harnesses the inherent advantages of the digital environment by realizing that if initial publishing costs can be covered through some means, the marginal cost of distribution can be almost zero.

The **Oberlin Group** seeks to develop its own open access eBook press through the Lever Initiative, while other initiatives pool together funds to pay publishers to release book content as open access. These are admirable initiatives, but likely to remain on the fringes of the academic publishing world until publishers find sustainable business models for them.

## Digital Interlibrary Loan?

Low-cost eBook packages, demand-driven acquisitions, consortial purchasing of eBooks, and open-access eBooks go a long way towards providing library patrons broad access to eBook content. Still, interlibrary lending of eBooks could play an important role in closing gaps in collections and providing broad opportunity of access to digital content.

The well-regarded academic book publisher **Springer** sells its eBooks to libraries in large, pricey packages without the option for à la carte purchase. The **Triangle Research Libraries** in North Carolina have struck a deal with **Springer** that allows some interlibrary loan of books between libraries in the consortium using library-created software that provides a limited-functionality copy of the book.

Local purchase with carefully limited interlibrary loan or for-fee-short-term-loan rights just may be the sweet spot that addresses the needs of library patrons while providing publishers a viable business model for digital books.

Libraries should not expect to dictate the terms of using and sharing eBooks as the Oberlin statement. Instead, libraries will need to negotiate sharing terms with publishers that work for both of them and

*continued on page 30*

# ATG Special Report — Big Data Takeaways

by **Ho Jung Yoo** (Research Data Curation Program, UC-San Diego Library) <hjyoo@ucsd.edu>

and **Reid Otsuji** (Research Data Curation Program, UC-San Diego Library) <rotsuji@ucsd.edu>

On March 12, 2014, **UC San Diego (UCSD)** held a two-hour special event called “Big Data at Work: A Conversation with the Experts” on the UCSD main campus. The purpose of the event was to showcase the topics of big data and data mining, where training in the latter is offered as a certificate program at **UCSD Extension**. The evening was kicked off with an introduction by **Hugo Villar Ph.D., MBA**, Director of Science and Technology at **UCSD Extension**.

What is this Big Data hype all about? What can we do with it? What is the future of Big Data? These questions were addressed by four experts representing the areas of academia, computing infrastructure, fraud and security, and commercial big data industry. The format of the event was 10-15 minute presentations followed by a panel discussion moderated by **Natasha Balac Ph. D.**, Director of the Center for Predictive Analytics at the **San Diego Supercomputer Center (SDSC)**. The four speakers were **Larry Smarr Ph.D.**, Founding Director of **CALIT2** and Professor of Computer Science & Engineering at **UCSD**; **Mike Norman Ph.D.**, Director of **SDSC**; **Stefan Savage Ph.D.**, Professor of Computer Science & Engineering; and **Michael Zeller Ph.D.**, CEO of **Zementis**, a San Diego-based software company focused on predictive analytics for big data and real-time scoring.

Several key points arose from the presentations. One is that the tsunami metaphor we often hear applied to the rise of big data is somewhat misleading. It's more of a sea-level rise or a high-speed elevator, because unlike the tsunami, which hits a peak then recedes, big data is only going to increase in size over time. Never before has humanity had to deal with this data growth rate. The growth phenomenon is coming from the multitudes of sensors (e.g., smartphones, automobiles, personal health devices) existing around the globe, their confluence, and the means of moving all those data around the network. A second point is that we often think of the three Vs that characterize big data: Volume, Velocity, and Variety. “Volume” refers to the vast amount of data being

accumulated and which has gone from being measured globally in terabytes to zettabytes. “Velocity” is the high speed at which big data needs to move; this movement is transitioning from batch movement to networked streaming data. “Variety” refers to the many different types of data out there created primarily by digital sensors. At a high level, data types can be classified as transitioning from structured data to both structured and unstructured data, which can be more challenging to deal with. But there is an additional fourth V that we need to consider when thinking of big data: adding “Value” to the other three Vs. In industry, the concern is business value, and in academia, value comes from information. To maximize the value of big data in industry, we need to move from descriptive analytics, which answers the question, “What happened?,” to predictive analytics, which asks, “What will happen next?” We also need to minimize latency from data to decision. This requires development of lots of automated models running simultaneously so that big data can be used to inform decisions in real time.

The speakers gave many examples of how we are deriving value from big data. **Zeller** noted the financial industry has been ahead of the game in terms of risk-scoring, prediction modeling, and fraud detection. Additionally, his key point was that big data in industries create opportunities to develop new platforms, capabilities, and business opportunities. **Savage** spoke of the security risks associated with having centralized data, and how **UCSD** researchers are using big data to improve security. For example, one group of researchers is using big data to battle email spam by behaving as naive spam users and conducting spam analytics. By analyzing those click trajectories, they were able to discover a great deal of spam traffic could be blocked by disabling just a handful of nodes (i.e., the spam enterprise or bank) within the spamming network.



In another example, the world's largest cryptocurrency, **Bitcoin**, had been touted for its pseudonymous transactions, which make transfers opaque. But are they truly opaque? Researchers at **UCSD** developed new data mining clustering techniques that allowed tracking of Bitcoin transactions in a way that made them less anonymous. The essence of big data security problems is understanding the information environment better, faster, and more efficiently than your adversary.

In the panel discussion, several interesting points were made about the usage and future of big data. **Zeller** called for the need to merge data analytics with the software technologies and to automate these processes so that predictive models based on big data can quickly inform our decisions. **Norman** described how **Jim Fowler**, a **UCSD** Professor of Political Science, has found there is a Facebook “Like” effect in national elections. Academic research is sometimes driven by events in the social sphere, and this leads research in unpredictable directions. When asked about the future of big data, **Zeller** predicts the hype will die down, but big data will have brought us lots of new, data-driven applications. **Norman** believes big data will give rise to a new discipline, much as “supercomputing,” once a buzz word, made way for the discipline of computational science. The development of educational programs in data science at universities will be driven by the public's desire for such training. **Smarr** predicts we will have intelligent personal assistants that will interact with us to inform us of our health, decisions, and interpersonal relationships. **Savage** foresees a major trend towards cloud computing. We will not be able to hold on to our own data anymore because they are being generated too quickly and are becoming too bulky and costly for transport. We will leave the data where they are, in the cloud, and send computational requests over high-speed networks to the data where they reside. 🌳

## Op Ed — Embracing the Digital ... from page 28

provide enough incentive for libraries and individuals to lease and buy eBooks on their own.

This is a healthy thing for the overall academic book marketplace. In fact, book-sharing between libraries could be viewed by presses as a sort of “freemium” approach to promoting their content: libraries and patrons that borrow virtual books offered by a publisher might be more likely to buy the book or other titles from the publisher outright.

Librarians are right to be concerned about the long-term integrity and stewardship of electronic book content. **Amazon's Orwellian** recall of 1984 is often invoked as a cautionary tale. But national or global preservation strategies by trusted players such as **Portico** or **HathiTrust** make more sense for long term preservation than relying on even the largest research libraries to create digital vaults for their eBooks.

Some readers still prefer print to eBooks, and the subject matter of some books just works best in print. Thus purchasing and managing print books will continue to remain an important, if

smaller, aspect of academic libraries' missions. Academic books are now commonly released in both digital and print formats, and for now, print collections and library sharing networks can serve as a fallback when digital access falters.

As books inevitably come to be released in digital-only format (see for example **Amazon Kindle Singles** in the consumer eBook world) libraries will have no choice but to come to terms with the the digital marketplace. Those libraries who swim with the digital current will have the most success in creating robust eBook collections for their patrons. 🌳