

November 2009

## Finding the Middle Way in Sustainable Digitization Efforts

Mitchell Davis

*BiblioLife*, [mitchell@bibliobazaar.com](mailto:mitchell@bibliobazaar.com)

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



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

---

### Recommended Citation

Davis, Mitchell (2009) "Finding the Middle Way in Sustainable Digitization Efforts," *Against the Grain*: Vol. 21: Iss. 5, Article 8.

DOI: <https://doi.org/10.7771/2380-176X.2644>

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.

# Finding the Middle Way in Sustainable Digitization Efforts

by **Mitchell Davis** (President of BiblioLife, Co-Founder Book Surge) <mitchell@bibliobazaar.com>



*“There will come a time when you believe everything is finished. That will be the beginning.” — Louis L’Amour*

At this year’s ALA in Chicago I had the chance to grab a coffee with **Jan Johnson**. **Jan** and her business partner **Michael Kerber** run the **Red Wheel Weiser** and **Konari Books** publishing imprints. Both imprints are leaders in the publication of esoteric, spiritual, and self-help texts. Our conversation focused primarily on commercial strategies to insure the ongoing availability of old and important books. In the realms where **Weiser** and **Konari** excel as publishers there is a strong theme of “re-discovery” — the old adage, *everything old is new again* has never been more pertinent than in the modern search for spiritual meaning.

As we discussed some possible projects and explored each other’s personal areas of interests we stumbled into discussing **Terrence McKenna**. **McKenna** was a gifted writer, speaker, academic, and cultural anthropologist who died in 2000 at the age of 53. Almost seven years after his passing, his personal library — collected from all corners of the world over the whole of his adult life — was destroyed in a fire. The fire broke out in a **Quiznos** near a storage area for the **Esalen Institute** in Berkeley, California. Many one-of-a-kind books carefully collected by one of the world’s best curators of such material — covering topics ranging from archaic philosophy, alchemy, shamanism and collective spirituality — were lost in the fire.

Does the **Esalen Institute** have content digitization at the top of its list of priorities? Probably not; they are busy running a retreat center just like most libraries are busy running libraries. But it is hard to argue the world is not a poorer place for these materials not having been digitized.

## Is Anything Still Out of Print?

These **ATG** articles were to contemplate: *is anything still out-of-print?* The question itself has a number of facets, but I have limited myself to addressing two of them in this article. Additionally, I was asked to provide some background context on print-on-demand — the technology without which most of this would be a moot conversation from a business standpoint.

As for answering the question about things being in or out-of-print:

- 1) The first part of the question speaks more to **packaging and different types of discovery**. It begs the question — *Does digital access mean something is “in-print”?* There are millions of books that exist digitally today, but for which you cannot easily purchase a reasonable print copy.
- 2) The second part of the question more literally speculates **do physical**

**books exist only in their original format that are not available for general public consumption (free or commercial)**. One could argue rather definitively that a number of books burned in that 2007 **Quiznos** fire are still (and possibly will forever remain) out-of-print. Tied closely to this is the challenge of how smaller organizations, associations and archives can financially support the important (but not prioritized) work of digitizing their unique materials in an era of declining public and private financial support.

## 15 Years of Print-On-Demand History in Three Bullet Points

I consider myself an un-official historian of the short and volatile history of print-on-demand. In a previous life (2000-2007) I was a founder of a company called **BookSurge**, which became the world’s first globally distributed POD retailing and manufacturing system. We entered that business very naïve as to the ways and means of publishing, printing and book distribution. We brought a beginner’s mind to the effort that helped us discover a new and better way of doing things from the pieces of existing systems. When we sold **BookSurge** to **Amazon.com** in 2005 we had built distributed POD facilities in seven countries around the world and had contracts to develop four more.

As I have been navigating the library world in my new role at **BiblioLife**, it has become clear that libraries have been thinking about POD for a long time. I wrote an article recently about Library / POD history (*Journal for Collaborative Librarianship* v.1, no. 3 - 2009), including the visionary *Project Ibid* launched way back in 1993 on through to groundbreaking projects by the **University of Michigan** and **Cornell** to commercialize digital materials in more recent times.

To oversimplify greatly, here are the three big developments of the last ten years that got us to where we are today:

- **Online Bookselling** — The Internet was born and soon after **Amazon** re-invented not only bookselling, but consumer behavior in general. **Amazon** removed the bounds of physical retailing and created a retailing platform that expanded infinitely the selection of media products available for purchase. We found out people will buy unexpected things if given the opportunity. The “Long Tail” was born and book publishing was re-invented. Where previously books selling less than 100 copies a year were a hassle, now it was a business. So were books that sold five

copies a year.

- **Self Publishing Accrued Successes** — in the late 90’s digital printing ushered in a new phenomenon as tens of thousands of authors began publishing books outside the scope and guise of traditional publishers. Some of them very successfully. The growth of this business helped push digital workflows and drove the print equipment vendors to produce more appropriate equipment for true (and profitable) one-book-at-a-time production. Traditional publishers owe a certain debt of gratitude for the current state of the art in POD to the self-publishing industry. Now, in 2009 you see large publishers considering using print-on-demand to develop “farm team” publishing systems. Some academic publishers do a majority of their print publishing in print-on-demand. But it all started in earnest with the self-pub houses.
- **The Marriage of Print-on-Demand to Channel Sales**. We can all thank **Ingram** and **Lightning Source** for this one. As I gave a presentation at ALA this past summer on the history of Print-on-Demand I related to the crowd that **LSI** launching in 1998 was the major reason we were all sitting in the room. While most printers were still looking at their publisher/customers as the progenitor of orders, **LSI** realized that it was the channel that mattered most to print-on-demand. By teaming up their technology with **Ingram’s** global channel they created new transactions and helped the technology sustain through the rough growing years. It has been fun to learn in the last couple of years how our launching **BookSurge** forced **LSI** to get better, faster and how that competition really pushed the infrastructure into something more usable for all publishers. Today, it all seems easy.

Over the past ten years the world has shrunk for those willing to think differently about how they manage a catalog and disseminate content. Libraries can be leaders in this new era of publishing on multiple fronts and of course we hope the work we are doing at **BiblioLife** will help facilitate that important paradigm shift.

With some context, we now explore....

## Does Digital Access Mean Something Is “In-print”?

As mentioned earlier in the article, there are millions of books that have been digitized that are not available in print today. If those books are available online (open or commercial) does that mean they are “in print”?

*continued on page 36*

As part and parcel of the **BiblioLife** digitization business model, we supply clean, post-processed, OCR files for our partners to mount on their own internal platforms at no charge. These files provide access to content in a managed way, but insure that all of the information can be accessed digitally for free. We know that it is of utmost importance to insure people have access to the collective knowledge held in the books being digitized. For libraries lacking a platform on which to host the content they digitize, we provide one at no cost. *However, as important as this access is to overall knowledge sharing, it is not the equivalent of a book being "in print."*

Print books must meet the qualitative expectations that have been set over the past millennia. Where book scans are of varying quality this is immeasurably harder than managing customer expectations giving away books on emerging electronic platforms. **BiblioLife** does human quality assurance and clean up on every book we publish. For free online viewing — hand prints, folded pages and other inevitable errors that occur in the digitization process are an acceptable experience. It is not, in our view, an acceptable experience to have a hand print in the middle of a page for a product costing \$15-\$25 and bearing the logo and brand of the contributing library. The difficult art of carefully completing the circle and making these materials available as reasonably priced, high quality print books takes huge steps toward insuring the posterity of the content.

When the **Wizard of Oz** was filmed no one contemplated DVD sales, but you can watch it on DVD today (you can buy a copy commercially or you can *borrow* it from a library). Commercial entities insure that the **Wizard of Oz** will continue to be available in any format that may emerge. The underpinnings of our eco-system are built on the same assumptions. We simply account for software automation to lower costs of making content commercially available and we maximize the sale of products on the long tail rather than the head.

### Moving from Cost Centers to Income Streams

With the basic infrastructure issues resolved for POD (and Digital) distribution, the challenge of bringing back out-of-print books becomes one of identification, digitization and post-processing economics, scalability, and product merchandising to interested customers (all of which could be articles unto themselves). These are difficult challenges for any resource constrained organization.

Book digitization as it exists today is typically a collection of cost centers for a library. Just the cost of non-destructive hardware to capture digital images has historically been a huge blocker to building effective digitization programs in all but the richest university and public libraries (a small portion of the library world in raw number terms).

There are also significant, but less visible

## against the grain people profile

Mitchell Davis

President, BiblioLife  
360 Concord, Suite 401, Charleston, SC  
<mitchell@bibliolife.com> [www.bibliolife.com](http://www.bibliolife.com)

**BORN AND LIVE:** Charleston, SC / Charleston, SC.

**EARLY LIFE:** Fun family life in small town, baseball, surfing, college, struggle to find self, tour guide, journalist, rock band, entrepreneur (in that order).

**PROFESSIONAL CAREER AND ACTIVITIES:** Founder and president of **BiblioLife** and **Organic Process** media/farm (an independent documentary, media production and local food co-op).

**FAMILY:** Wife **Farrah**, two beagles (**Freddy** and **Lucy**), Cat (**Natalie**). Amazing parents and brother.

**IN MY SPARE TIME:** **Farrah**, documentaries, travel, music.

**FAVORITE BOOKS:** (Hard to narrow this down) — *An Intimate History of Humanity* by **Theodore Zeldin**, *Jitterbug Perfume* by **Tom Robbins**, *Archaic Revival* by **Terrence McKenna**, *The Wisdom of Insecurity* by **Alan Watts**, *Inside Nothing* by **Ben Timpson**, *Trout Fishing in America* by **Richard Brautigan**.

**PET PEEVES:** Over-thinking, over-talking, under-doing.

**MOST MEMORABLE CAREER ACHIEVEMENT:** Starting, growing, selling and successfully integrating **BookSurge** into **Amazon.com**.

**GOAL I HOPE TO ACHIEVE FIVE YEARS FROM NOW:** To have contributed to the inevitable obliteration of the line between business and philanthropic pursuit.

**HOW/WHERE DO I SEE THE INDUSTRY IN FIVE YEARS:** Hard to predict — but certainly more books in more places and a book supply chain that more than ever before — rewards those who add value and is increasingly painful for those that do not. 🐻

infrastructure costs of actually handling those digital images in a way that provides ongoing value. This includes image post-processing, OCR, discoverability, packaging and the ability to deliver in multiple formats, etc. Most digitization projects rely on internal funding from operational budgets (where it is typically difficult to fund digitization at any scale whatsoever) or from one-time grants (also rare relative to the number of libraries needing to digitize unique content). The latter approach can leave projects half completed or unable to be supported over the long term.

**BiblioLife** has never asked a library for money. Our mission from the start (and we have been very successful thus far) was to create new dollars from new demographic customer sets. We lower the barrier of entry significantly to participating in digitization and then extract content, expertise and energy from the library in a working relationship. This subtle shift away from being a "vendor" trying to sell products to libraries is actually tectonic for unlocking the potential of libraries as curators and distributors of content.

A library dean participating in the **Shelf2Life** program (a digitization program that we operate in partnership with the **Bibliographic Center for Research**) summed it up in saying:

*"As libraries we are grounded in the culture of the book. We need to find ways to bring the culture of the book into the future and for us to think as libraries about how that cultural value — buy it once and make it available time and time*

*again — can translate into something that helps us do what we do...revenue... is a new and different idea."*

At the end of the day **BiblioLife** is in the artifact business and we get up every day figuring out how to share more artifacts, with more people, in an economically sustainable way.

### We Have Barely Scratched the Surface

To the question — *do physical books exist only in their original publication format that are not available for general public consumption (free or commercial)*, it should be clear that beyond the small handful of major research and national libraries lie tens of thousands of collections similar to the **McKenna** collection. It is estimated there are 60-90 MM out-of-print books available for digitization globally (some academics place the number as high as 150 MM). Whether looking within the archives of historic societies, associations or small public libraries — the work of digitizing the world's unique books has just begun. And the work of quantitatively and qualitatively organizing those books is even further behind.

Similarly there are orders of magnitude more public domain content in non Latin-1 formats, which can only be discovered by using an ASCII phonetic interpretation of the word. Accordingly, the UTF-8 standard, a variable length character set, was introduced that addresses these languages. Unfortunately, most common Web stores such as **Amazon.com**

*continued on page 38*

# It Is Very Much an “E” and “POD” World

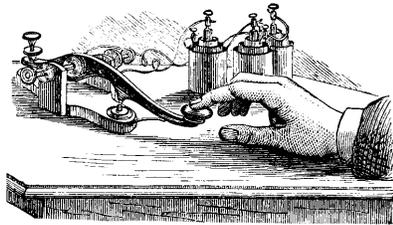
by **David Taylor** (Senior Vice President, Content Acquisition, Ingram Content Group, President, Lightning Source)

Having been involved in the bookselling business since the mid 80s I have done my share of the world's book fairs. **Frankfurt** looms large in my experience and not just because of the strangely unappealing German food, or the moans from the large British contingent about the price of hotels or the impossibility of getting a decent cup of tea.

The thing that always sticks in my head are the endless rounds of seminars, roundtables and presentations where the collective minds of the world's book trade denizens get up on their hind legs, switch on the **PowerPoint** slides and attempt to predict what is going to happen to the book, and the trade that revolves around it.

I know. I am one of those denizens and I too have done my fair share of predicting and more than my fair share of getting it badly wrong. I may be about to do it again in what follows but you and time can be the judges of that.

Anyone who attended the **Frankfurt Book Fair** around the year 2000 will recall the intense and almost feverish excitement that accompanied what might be called the first eBook wave. Disciples of the new technology vied with each other to predict the death of the physical book or “tree ware” as some wag coined paper books. There was also the most fantastically lavish eBook awards ceremony that lasted I think for about three or maybe four years. Lashings of drinkables and plates of clearly very expensive nibbles as the Brits call them, were eagerly consumed before the devotees were ushered into the awards ceremony where a succession of unrecognized authors were awarded prizes for the best eBook fiction and so on. The worth of the content subsumed in the format of its delivery. All very odd I thought. The event petered out after it became clear to all that the expected promised land of a paper book free world was simply not going to turn up and



the **Frankfurt** trenchermen went in search of alternative fodder.

As 2009 draws to a close, eBooks are coming to the fore again, powered by improved technologies, the introduction of devices like the **Kindle** and the **Sony** reader and a more sober and informed understanding by both publishers and booksellers about how content should be packaged for consumers who want to read it. The simultaneous release in both the e-book and print format of **Dan Brown's** *The Lost Symbol* will probably be looked back on as a key event in this trend, for consumers are being offered that oldest of options: choice. This is all a little different from the early insinuations of the eBook pioneers that consumers should be weaned from their odd notions of wanting to have a paper book. Consumers tend to make up their own minds.

The book trade world of 2009 is not a world that is seeing eBooks as a replacement technology for the vast majority of content that is currently in a physical format, although there are clearly areas of publishing that fit the eBook format better than others such as high end academic and scientific titles. The world of 2009 is one in which content, the stuff that authors produce, is increasingly being offered into the market in both e format and paper format. The major shift however is that the technology of print-on-demand is moving to the very centre of how a publisher offers content in the traditional physical paper form. This shift is happening hand in glove with increasing interest in the offering of content in a variety of different ways. “Either and” has replaced “either or.” This is not an audio cassette being replaced by a CD moment or CD's being replaced by downloads.

I have been involved in the world of print-on-demand (POD) at **Lightning Source** for a little over six years. In that short time, the technology of digital printing

has leapt forward. POD produced books are no longer derided as inferior quality. POD correctly defined is the ability to produce a single copy of a book from a digital file; it is not, despite the best efforts of some digital printers, about ultra short run digital printing because that is still premised upon the production and holding of speculative inventory in the hope that it will be sold. Pure POD allows a book to be sold and then printed, and that has truly radical implications for the book business if you start thinking about it on a global scale.

There is another factor that has played here though. The welding of scalable single copy production capability to large scale global distribution via wholesalers, distributors and Internet booksellers has the potential to effect massive change on the book publishing business. This is the model of supply that has been pioneered by **Ingram** via the marrying up of **Lightning Source** with its wholesaling and book distribution capabilities. It is this combination that is driving some very radical thinking in some of the world's largest publishers and distributors.

The advantages of POD for a traditional publishing business are very well understood: the ability to reduce inventory, to keep books and associated revenue alive, to reduce the risks of over printing and excess inventory and so on. POD has also given rise to whole new types of publishing models from self-publishing, to the resurrection of public domain books that have long been unavailable to buy.

What we are now starting to see is an awakening realization amongst publishers, partly fuelled by the economic downturn that this marriage between POD and global distribution is going to allow a fundamental shift in the very way that they get their content into global markets. Further, the improvements in digital printing and finishing, and the promise of further shifts to come with the advent of digital ink based rather than toner based printing are going to allow even more of their titles to come within the remit of the POD production model. Now we are not just talking about POD allowing long tail titles to remain in print or authors to self-publish; we are talking about a shift from a global book supply chain that is based on warehouses filled with speculative inventory to one based on minimal inventory or no inventory at all. The distribution centre full of books printed speculatively is going to be replaced by much smaller footprint centres only holding those titles in a preprinted format that POD cannot currently manufacture. Increasing numbers of titles are going to be printed only when a sale is made. In addition, the idea of printing books in one location in a large print run and then shipping them around the world to sit in regional warehouses waiting for an order is becoming a defunct one. **Lightning Source** has already demonstrated the attractiveness of this model of production and distribution via

*continued on page 40*

## Finding the Middle Way from page 36

have been slow to support UTF-8 because it could make the user experience confusing, and most customers don't require it or even have keyboards to support it. The result has been that almost all Web retailers are not global but regional, and there remains no effective and central place for non-Latin-1 language products.

### The Middle Way

The work to digitize books in the nooks

and crannies, while remaining economically sustainable, requires more patience than a big box strategy funded to the tune of hundreds of millions of dollars. It requires a more nimble infrastructure, innovation in processes to lower the cost basis of putting up books for sale, and a keen sensitivity to doing win-win deals on a global scale. There are some important forces at play and they require a different type of thinking. Our approach is to build a practical “middle-way” that realizes a task of this magnitude will take time and that places libraries at the center of a “Local Media” movement that can contribute to a global knowledge eco-system. 🌱