People Profile: Mitchell Davis

Editor
Finding the Middle Way
from page 34

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 posternity 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 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...