

Purdue University
Purdue e-Pubs

Charleston Library Conference

The Evolution of E-Books

Mitchell Davis
BiblioLabs

David Durant
East Carolina University

James O'Donnell
Arizona State University

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

An indexed, print copy of the Proceedings is also available for purchase at:

<http://www.thepress.purdue.edu/series/charleston>.

You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: <http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences>.

Mitchell Davis, David Durant, and James O'Donnell, "The Evolution of E-Books" (2016). *Proceedings of the Charleston Library Conference*.

<http://dx.doi.org/10.5703/1288284316496>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

The Evolution of E-Books

Mitchell Davis, Founder & CEO, BiblioLabs

David Durant, Collection Development Librarian, East Carolina University

James O'Donnell, University Librarian, Arizona State University

The following is a transcript of a live presentation from the 2016 Charleston Conference.

James O'Donnell: Thank you, Tony, and good afternoon. My view of e-Books resembles Gandhi's view of Western civilization. If you ask me what I think of e-books, I'm likely to say, being only slightly provocative, "Sounds like a good idea, and I hope somebody invents something like that someday." Okay, I know I'm being provocative. On other days, I might just easily like to claim that I published the first scholarly monograph ever distributed over the Internet, a book about Dante's Epistle to Cangrande which we distributed by Gopher back in 1994 from the University of Michigan Press. We have been not inventing the e-book for a good long time, but the challenge of the things we now call "e-books" is that they aren't books, and they are only moderately "e," and they just plain don't work very well. I could speak to the mass market product, the ePub form, the Kindle form, which more resembles the papyrus roll of antiquity than anything else I know. As long as you're willing to start on page 1 and scroll through to the end, they work pretty well. They are good for reading a murder mystery. Get to footnote number one, and you're in trouble already. Try to look at map number three, and you are deep in trouble, and if you want to see the illustrations, the chances are they will be small and blurry and unintelligible if they are there at all. Those things can do something. They can't do everything. At their very best, and I think the Kindle e-book is probably the e-book of the moment at its very best, they are considerably dysfunctional, and we could have a show of hands of how many of you believe that the Kindle format and the Kindle formatted e-books will be here 50 or 100 years from now in a reliably preserved form that you can look at. No, no chance of that. No chance of that whatsoever. Even the tackiest, old, cheap paperback from the 1950s had a better chance of lasting to 2016 than the Kindle books we have today. I say at their best they are dysfunctional. At their worst, they are deliberately crippled by the people who sell them so as to sustain the business models according

to which they sell them. Go to one of the expensively purchased e-books in the Arizona State University Library catalog, click through to it, stop if it happens to come from vendor (mumbles, covering mouth), and fill out a new form for a new login and password so that you can have an account with the vendor, even though you've authenticated your way in to it through the ASU system. When you get there, you will discover that it is a 400-page book, and you are entitled to download 60 pages of it, or print 30 pages of it, or maybe it's the other way around. I always get confused. But you could download it to your reader for 14 days, thereby depriving any other ASU user of seeing it for those 14 days, and with luck it, will expire on your device after 14 days. If you've used it for 10 minutes, you have no incentive to give it back, so another 13 days, 23 hours and 50 minutes must spin by before any other user can come to them. They don't do footnotes any better than the Kindle does. Their dealing with graphics is sketchy at best. They're probably deriving from publisher-supplied PDF files, and publisher-supplied PDF files are about as good and about as bad as you might expect them to be.

To make the point that they are dysfunctional and crippled, I will call your attention to the competitor that has come into existence in the last year or so, four of those legally purchased and expensively paid for e-books coming through the Arizona State University library. I mean Sci-Hub and LibGen, brackets: I make no apologies for the illegality or the sketchiness of Sci-Hub and LibGen. In fact, I have suspicions that there is a dark side to that enterprise that has not yet achieved very wide attention. Let me just say that if I were invited to hand over my credentials to Sci-Hub or LibGen so they could download articles from my university server, because I am a noble person and believe in open access and the liberation of information, I would know that I should also be expecting those credentials to be scrutinized by a bunch of guys in a back room in Moscow someplace with a lot of tattoos, a lot of piercings, and a lot of cigarette

smoke around them and not the highest moral standards that I would want participants in the scholarly publishing system to live up to. End of bracket. But, if I am given a choice at this moment between the e-books that we pay expensive prices for, I believe I personally caused ASU to spend \$195 a couple of weeks ago because I needed to print 30 pages from one e-book, and it didn't work the first time around, so I went back around again to do it again, fully aware that we have a deal with that vendor for two uses equals purchase, so we had triggered the purchase price. I do that all the time. If I have a choice between one of those and a bootleg PDF provided to me by the Sci-Hub, LibGen interface, I believe that if I am a rational person I will prefer the bootleg product. This is nuts. In the great world of commercial products, if you go out to buy a bootleg Louis Vuitton handbag, you will spend about \$30 for it, and you know as part of the contractual deal of purchase that it will fall apart on the third use. That's how it goes. But bootleg PDF's from Sci-Hub LibGen don't fall apart on the third use. In fact, they are whole lot easier. You can print all of the pages that you want to print. You can read them any time. You can carry them around with you where there is no network on your device. You can cut and paste. You can—the better quality your Acrobat reader is, you can extract text from them. You can do all the things you only dream of doing on the expensive product that we are paying for. Again, I say I do not defend the people. I do not defend the system that gets us there, but I point to those facts about bootleg content as a way of demonstrating just how crippled we are and where my initial provocation about hoping that e-books get invented someday makes some sense because of the absurdities that we have inherited.

So, I'm going to say concisely four things that I think we need in order to get to a point where we think we've invented the e-book. First, we need a lot fewer platforms and better standards. When a vendor comes to me and says, "Jim, I know it's a problem, but we are rolling out a new platform next year." I feel exactly as I feel when Microsoft tells me that there is another version of Microsoft Word coming, Office 999, or whatever it might be, that will be bloated with new features that I don't know how to use, and all the buttons will have migrated someplace else, and I'm going to have to figure it out if I'm going to stay in that space. Navigating from Clem Kadiddlehopper's e-book platform to Ralph Kramden's e-book platform to Ricky Ricardo's e-book

platform is no damn fun at all. That's one of Sci-Hub's advantage. They haven't got a platform. They've got a simple, federated search for everything they've got, and the interface for everything they've got is pretty close to consistent. Let's see, if I remember from them the one lingering problem is that sometimes you got EPUB, sometimes you've got PDFs. Sometimes you've got DjVu. Sometimes you've got MOBI, and occasionally we even see that Microsoft oldie goldie. How many of you are old enough to remember the LIT format that was going to solve all these problems for us? We need to get past individual platforms. We need to get to the most open possible standards, the most consistent possible standards and, therefore, the greatest degree of discoverability and usability for the items that we use.

We need more functionality. I've got to be able to print the whole thing. I've got to be able to take the whole thing with me. I've got to be able to do the things that I want to do with the technological possibilities of the format you are presenting me without being held up, stopped, frisked by border guards, told to move along now, told I'm not allowed to take pictures here, whatever else that makes it feel like we're living in the national security state.

I will make one more suggestion in two parts, that is that we probably need a solution for the e-book that is designed to suit the print book in e-form, mainly therefore a legacy service. That is to say that it can take as many as possible of the features that generations and generations of compositors and printers have devised from making the printed book so functional and provide useful equivalents for them. I've got to be able to see the pictures. I've got to be able to read the legends on the map. I've got to be able to get back and forth to the notes quickly and easily. I hate endnotes in printed books, but at least I have a thumb and forefinger and know how to flip back and forth. This technology is better than what any of our e-book vendors are providing us now for flipping back and forth to the endnotes at the back of one of their books. We, therefore, probably need another solution for the e-book going forward. The packaging of the sustained argument, the sustained narrative, it's associated appurtenances that we now can see in digital form and for that matter, if it is a scholarly book, all the data sets the author wants to bring together with some opportunity to print something but not an opportunity only to imagine the print artifact as the

original version but to imagine print as one user interface that can make sense with others. I think there are exciting possibilities there. Maybe there's somebody at this meeting already who knows something like that that's happening. I'd love to hear it. I'm easy to find: jod@asu.edu. I think that's the real future for somebody who will make a ton and three quarters of money selling e-books and will leave even me convinced that we have finally discovered what e-books might be. As for Western civilization, talk to me after Tuesday. Thank you.

David Durant: Thank you and good afternoon. I have been left with a lot to live up to here, so I will do what I can. My remarks are not going to be so much about the actual technology, the e-book and about e-reading technology per se as about the broader impact they've had on our society, the possible future impact, and what we as librarians, especially as academic librarians, should take into account as we organize our own collections across all formats and future.

At the beginning of this century, e-books were, to put it bluntly, in terms of their popular appeal, there were something of a damp squib. You had wonderful projects like Project Gutenberg, and in about 2003 or 2004, I believe you had NetLibrary, which we could subscribe to and add e-books to our collection, but it wasn't really until 2007 when the Kindle was invented that the e-book really exploded, that it took off. And so, for example, by 2011 Amazon announced that its e-book sales had actually exceeded its sales of print books. By 2012, a survey of American publishers revealed that their sales of e-books had gone from 10 million in 2008 to 457 million in 2012. So, this sort of explosion of popularity of e-books in a very, very short period of time led to what I would call the "substitution model." You had print books. Now you have e-books. We're going to get rid of the print books and replace them with e-books, essentially, to simplify it greatly. Essentially what happened in terms of print journals—substituting electronic journals for print journals. Substituting print reference books for electronic reference items and electronic reference databases, so this belief tended to be quite popular by about four or five years ago especially.

However, a funny thing happened on the way to our "all digital" reading future. There were voices, first of all by popular writers such as Nicholas Carr especially, as well as academics like Maryanne Wolf and Naomi Baron, a lot of concern about the impact

that electronic reading and reading off e-devices as opposed to reading off of the printed page was having on our ability to engage in reading and to engage in particular in what is called linear or long-form reading, the ability to read at length, in depth, for a considerable period of time and then to be able to memorize and absorb that material and incorporate it into our pre-existing base of knowledge. There's a lot of evidence, not just anecdotal but scholarly in many cases, that reading off of most digital devices tends to hinder, it encourages what is called tabular reading, reading short bits of information for sort of a quick piece of information here and there as opposed to being able to read at length and in-depth, with potentially great impact on our society. In addition, the spread of e-reading and e-readers in particular, e-books has sort of plateaued in the last several years. By 2014, e-books had risen to about 30% of major publisher sales here in the US. Since then, they've kind of leveled off, and even I believe, according to some figures, they've actually dipped slightly. At the same time, print book stores last year recorded an increase in sales of print items of 2.5% here in the United States. That was the first increase in print bookstore sales since 2007. So there is a sense, a growing sense that e-books, that e-reading technology, the popularity of e-books has sort of leveled off. In particular, one phenomenon worthy of note is that e-readers, dedicated e-readers, Kindles, Nooks, their popularity has leveled off. Nearly 20 million e-readers were sold worldwide in the year 2011, according to Forrester Research. By 2014, that figure had declined to about 12 million. So, e-readers sales have declined to the point where many people are speculating that the e-readers, essentially, the dedicated e-reader, that Nook, the Kindle is headed for a sort of boutique status like the digital camera, and people are increasingly reading off of their smartphones, their tablets.

For example, the 2016 Pew Reader Survey came out just in September, has found that only 8% of people who read a book in electronic format read it off of a dedicated e-reader device. 15% read at least one e-book off of a tablet. Going back to that Pew study, the Pew studies have been done on an annual basis, and so they found some fairly consistent results that people are not generally speaking abandoning print for reading e-books. For example, the 2016 study found that 73% of people are reading books, had read at least one book in at least print or electronic format, of the total number of respondents 65%

read a book in print. 28% read at least one electronic book, and those numbers have been fairly steady. The number of people reading a least one e-book per year according to Pew has essentially been either 28 or 27% over the last three years. The number of people reading a least one book in print format, according to Pew, has essentially fluctuated from 63% to 69% to 65% this year, according to Pew's figures. So, that figure is staying roughly the same ballpark. Generally speaking, the statistics we have show in user preferences, in media articles you see a lot of articles about millennials actually prefer to read print for long form. This is reflected in statistics, so there seems to be a general sense that, at least for now, that e-reading, e-books have sort of found their level, and that what this leads me to conclude, and leads many others to conclude, is that we should not think of a substitution model in terms of e-books versus print. We should think of a complementary model in that reading off the printed page and reading off a digital e-reading device don't necessarily facilitate the same types of reading. They each have their uses. Each of them has their own uses, and us as academic librarians, as we build our collections and future, we need to recognize the differences between reading off the printed page and the preferences of our users in terms of reading off the printed page versus reading off a digital screen. Understand that each of these facilitates a different form of reading, generally speaking, and build our collections so that we offer our readers the best of all possible worlds in terms of reading. The ability to engage in in-depth, lengthy, linear reading off the printed page as they need to, and also the ability to engage in tabular reading, searching across text, finding brief bits of information or short passages on a particular topic off of the digital reading devices. And so, we need to facilitate both these forms of reading as we go forward. That doesn't mean that our print collections need to stay the same size they are now. It doesn't mean we can't weed print materials or that we have to continue buying as many print materials as we do now, or that we can't send print materials, at least some of them, to our remote storage facilities, but we need to make sure that we sort of retain an echo system of reading in which print and digital are coordinated, are sort of integrated to best meet the broad spectrum of the needs of our users.

In short, we need to facilitate, support what scholar Maryanne Wolf has called the biliterate brain: An ability that can both read linear, in-depth text, as

well as engaging in the short tabular reading, but while it can be done in print and has been, is especially facilitated by the digital reading environment. Thank you.

Mitchell Davis: So, thank you guys all for coming out. I know the porches and decks of Charleston happy hours are calling everyone, so I appreciate you guys being here so late. I think I've got a little, and I think this is a great mix of people to talk about this because I'm the only one up here who isn't a librarian and isn't from an academic library. We're a software company, and we license software and license content to libraries, and we work across a number of different markets, so we work in K-12 libraries. Most of our attention over the last two years has been in public libraries, which face a similar challenge as academic libraries, I think, but in a much less insulated world, and so public libraries are really having to compete toe to toe with the Amazons of the world and the media companies of the world that are just making enormous amounts of content available for very small amounts of money and really challenging what the role of the library is in a digital future. So, I bring a couple of different perspectives to that.

I've been coming to the Charleston Conference since 2001. I was around in that early round of reciprocal and NetLibrary and all those e-book aggregators, some of which were consolidated, some of which disappeared overnight, and did all sorts of different things. I think we really nailed it with Amazon. That's what I was going to say is that all attempts at e-books really just had been playing at it until Amazon launched the Kindle. And even though that first device was ugly and clunky, they got it out, and the one thing they got right was it knew how to take your money and give you a book, which Amazon tends to always get right. So, and they've of course made the device enormously better as it has moved forward.

I talked yesterday, and I think that the past of the e-books has sort of gotten us to here, and we talked about a lot of the problems with the price of the e-books and the functionality of e-books, and for me as a startup software company in the library industry, and if any of you guys did not know, this is a harrowing industry to start a software company. Not sure if you knew that or not, but it is. It really speaks to the business structures. Most of the companies that are selling these books and these technologies are private equity-run companies.

Innovation is anathema to the entire mission of private equity. Don't know if you knew that, but that is also true. Their job is to create consistent returns to investors, and innovation and disruption actually is a threat to that. So as long as those companies, the checks are being written to those companies, and small innovative companies are forced out of business or forced to consolidate, it is very hard to imagine how the future improves, honestly. The thing that I think is looming that I hope can sort of coalesce effort and coalesce attention is that Amazon is definitely coming for this industry, and maybe that excites universities. It definitely doesn't excite vendors. I know that. But, it may excite universities. Who knows? I'm not making a judgment on it, but with their open education, with their OER effort, Amazon and SPIRE, make no mistake that they can suck all the profit out of this business. They could care less. They don't care about library companies pushing each other around for market share. They will just come in and reinvent the whole thing. I think the universities spending the money really have to decide is that the future we want, or how are we going to fix this, and certainly making it easier for innovative companies to succeed I think is part and parcel of that.

One of the things that we've been thinking about is as we've sort of—we have a big project that we've been doing in the UK with JISC for the last two years. It's an open education textbook project. And in that project, I think the thing that I found is that here in the US of course of the last 15 years, 20 years, the price of education has skyrocketed, the management class of universities has expanded, and it really has become more of a business. Kids are still learning things. Nobody is going to stop that from happening, but the machine of universities is a business. And I think what I see coming and what I think is good news for everyone is that in an effort to enhance student user experience, to make sure that the people writing those checks feel good about writing those checks, that the parents are happy, that the students are happy, that textbooks, overpriced textbooks, are the first thing that the university management is going to put in their crosshairs because they can improve the user experience without taking one penny out of their own pockets. So, I think that OER is going to be able to usher that in. I think that once there is enough critical mass of success with OER textbooks, I think university administration is going to get it immediately, that we can save students \$2,000 a year and not lose any

money. We are not selling any of these books out of our university bookstore now anyways; they're all being ordered from Amazon and Chegg, so we don't lose anything there. And it's going to be a pretty radical change, and you see it coming. Because in the UK, education was free until a few years ago, and it just immediately went to 9,000 pounds a year. Most of the projects in the UK, most of the studies being done in the UK are focused on student experience because they have to justify going from free to 9,000 pounds a year. But, here in the US, it's been more like the frog being boiled in water, you know. It's just kind of slowly happened, and so you don't see as much focus on that.

And so, we just got done with our first semester pilot at Liverpool. The feedback from the students was phenomenal because that's really all we focus on. We focus on the student user experience and delivering all media types through one single interface so that students aren't jumping out to watch a video there, an image here, a database here, and an e-book there. It all happens in the same interface. Everything is unlimited simultaneous use. There are no checkouts, returns, holds, turn-aways, so it is the kind of experience people have in their real life being brought into the university, which I think is what students expect. We've been very happy with that, so we are moving pretty aggressively into that.

I think the other thing that I would say just to temper a couple of things that were said before me, is that a lot of the data, and just think about this critically as you see more and more e-book data come out, and this probably relates more to the trade market than the academic market, but Amazon doesn't give data to anyone, so any data analysis you see on e-book usage coming from traditional publishers doesn't include 95% of the pertinent data. So, if a traditional publisher says e-book sales are flat or going down, what that means is more people are reading on Amazon, and no one can see that data. They are reading on Kindle Online Lending Library. They're buying books directly from Amazon Publishing Imprints. One of the most fascinating sites, if you want to keep up with that e-book market, is a site called authorearnings.com where this guy has written a bot that crawls the public Amazon website and looks for changes in sales ranks and all different sorts of things. He believes he has about 85% coverage on the transactions that happen on Amazon, which is a

phenomenal accomplishment, and he every three months publishes the data on author earnings, which I also think is a massive shift, right? He's not measuring the health of the publishing industry on publisher profits or stock price. He's managing the health of the industry on author earnings, which is completely agnostic as to how you published your book, whether you self-published it, traditional publisher, small publisher; it's really phenomenal. And what he is finding, it swings wildly from one three months to another, but the last report he put out was that big five author earnings in the last two years, and this is again trade, but have been cut in half, and self-published earnings have doubled in the

last two years. So, there is clearly this indie revolution happening in trade, and I see a lot of parallels in that with OER that if you've got the right curation systems in place, if you've got the right peer-review, if you've got the right abilities to sort of push books where they need to go, that this indie revolution in a curated way can really replace these \$200 textbooks pretty quickly. So, we're very excited about OER. Our office is also right across the street, 100 Calhoun, and we've got an open house the next two days, so if any of you guys want to stop by and see what we're doing with OER, we would love to have you over there. Thank you.