# E-books
# One librarian's perspective

James Mouw
Director of Technical and Electronic Services
The University of Chicago Library
mouw@uchicago.edu

THE UNIVERSITY OF CHICAGO Library

I would like to thank the STM and the program planners for including me on this program, it is a great pleasure to be here.

Introduce me.

I'd like to begin with just a few words about Chicago and why we're not typical.

>Big
>Private
>Comprehensive = exhaustive
>Intensely academic = broad and deep collections and access with heavy use of both print and online collections
>Faculty in Social Sciences and Humanities view the library as their laboratory and are both supportive and protective.

Indeed, we're building new on-campus stack space

# The Joe and Rika Mansueto Library



…since we're still adding as many print volumes ~180,000 per year- as always. This facility will hold 3.5 million print volumes and is expected to be filled in about 20 years at our projected rate of growth.

I also find that we're worrying about very different things than we were a few years ago, and they seem to be the same things publishers are worrying about:

        transforming our technical infrastructure
        the importance of branding and communicating our mission
        and, we are also deeply into the digital age in our own arena, including a move to 100% digital dissertations, and we're struggling with
        all the same things publishers are

# Agenda

- A brief look back at how far we've come
- How we spend our budget
- What our users are telling us
- Not all e-books are created equal
- How they are being used
- How this affects print purchasing and usage
- Role of various readers/delivery platforms
- The importance of access mechanisms
- Interesting new partnerships developing
- ROI – we need data data data data data

# Remember Serials?

Sample Online Serial Full-Text Downloads

2009 Jan-Jun actual with projections

| Product | # of Titles | Downloads Jan-June | Jan-June per title | Proj. 2009 per title |
|---|---|---|---|---|
| JSTOR | 1,027 | 385,448 | 375 | 751 |
| Elsevier | 2,470 | 765,812 | 310 | 620 |
| Wiley | 1,898 | 121,057 | 64 | 128 |
| Nature | 93 | 138,471 | 1,489 | 2,978 |
| Science | 1 | 34,341 | 34,341 | 68,682 |
| Springer | 2,232 | 67,974 | 30 | 61 |
| ACS | 44 | 102,127 | 2,321 | 4,642 |
| Oxford | 229 | 40,272 | 176 | 352 |

**All data pulled from Scholarly Statistics, 25/09/2009

Talk about the process of conversion from P to E serials, immediate acceptance and why that might be. Why books are different. Emphasize that in all cases I've pulled a very small number of sample stats from the hundreds of examples I could have used, meant to illustrate particular common types of situations, not comprehensive.

# JSTOR as example – Return on Investment

| Year | Dollars | Articles Viewed | Cost per View |
|---|---|---|---|
| 1997 | $3,750 | 6,105 | $0.61 |
| 1999 | $3,750 | 44,528 | $0.08 |
| 2001 | $11,250 | 80,845 | $0.41 |
| 2003 | $26,286 | 237,964 | $0.11 |
| 2005 | $41,538 | 360,936 | $0.12 |
| 2007 | $37,965 | 874,012 | $0.04 |
| 2008 | $54,225 | 796,361 | $0.06 |
| 2009 (projected) | $56,625 | 770,896 | $0.07 |

And serials provided us with an very good ROI, with little or no substitution for print use, mainly new uses, at a very high level

# Budget allocation
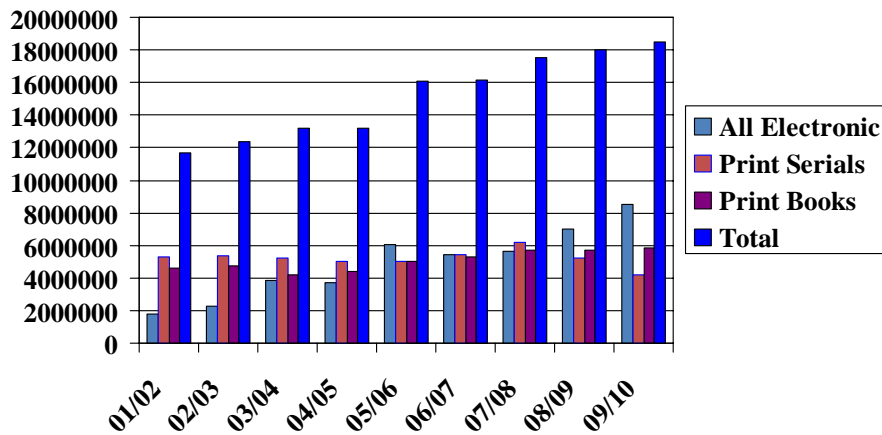
A quick look at how my library is spending its collection money.

The trend we see is that the budget has always been up – largely because we've protected collections at the expense of other things. While the budget is still up slightly for FY10 the increases came in entailed endowments not the general allocations, which were actually reduced. This trend is likely to continue for FY11 and FY12.

You will note that for my library there are three basic trends

Print serial purchasing is on a steady decline
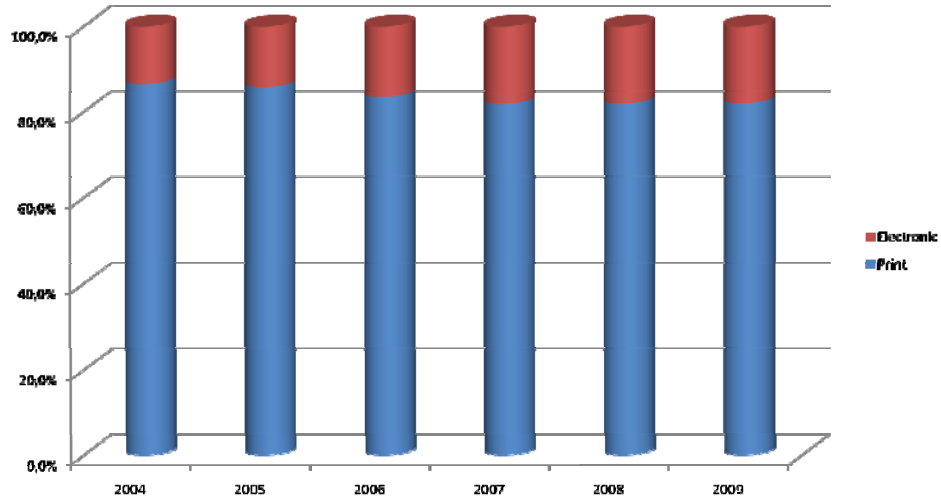Electronic purchasing is way up, and with a big jump in the past two years
Print monographs are holding steady, with virtually no change in the past decade

Increasingly, these CATEGORIES ARE NOT CLEAR as it becomes harder and harder to budget print separately from electronic

Titles added – year of publication
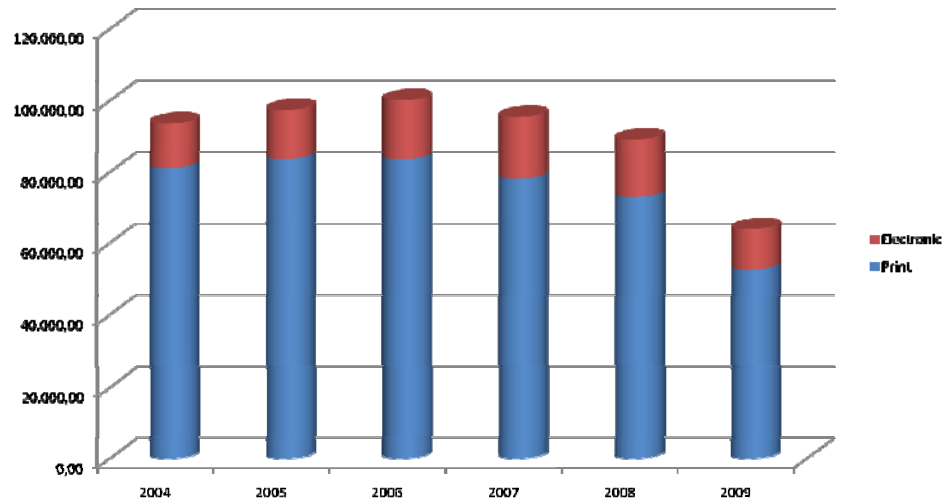
# Titles added by Percent (Year of Pub)

E-access gradually rising from 11% to almost 18% (one-fifth). Actually a much smaller percentage than some libraries are reporting, especially where large chunks of their budgets are being turned over to patron-driven purchase.

# Patron acceptance

- Widespread acceptance of Serials
- Very uncertain about whether the same would hold for e-books
- Early experiments received mediocre acceptance
- Originally e-books seen by libraries as additional means of access not instead of print

# Gartner Hype Cycle
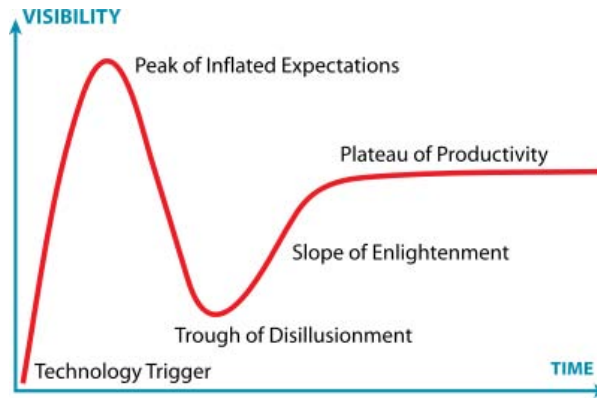
In retrospect, the process of conversion from P to E serials was relatively painless with immediate acceptance for all sorts of reasons, primarily the fact that almost no one reads an STM journal from cover to cover, and that the articles were a prefect size for easy electronic delivery. Discoverability structure already electronic and in common use. Serials jumped right to "peak of inflated expectations" and STAYED THERE.

Books never hit the initial bump and went directly to the "trough of disillusionment" only now on the "slope of enlightenment"
Books are different in several significant ways:
          many are stand-alone treatises – patrons want to read an entire book, and authors don't want their baby "chopped into pieces"
          much lower usage rate than journals
          Print books were (and continue to be) largely individual purchase decisions
          They have their own unique selection, budgeting and purchasing stream
          Initially, while serials were largely an "instead of" decision relatively quickly, for BOOKS at this point it remains largely an "in addition to" proposition

# Types of e-books

- Scholarly Monographs – OSO
- Mass market – Netlibrary
- STM – Springer
- Technical – Safari
- Major reference works – Credo, ORO, Gale
- Retrospective collections – Serial Set, EECO, EEBO, MOML, MOME
- Textbooks

Intro slide to the following specific slides

# Scholarly Monographs
### Oxford Scholarship Online

- Electronic version
  - 2,765 titles currently purchased
  - ~20,000 uses per year
  - 17 uses per title on average per year
- Print collection of same titles
  - 2,381 titles in collection
  - .43 average circulations per title per year – 10 titles more than 50
  - 3.81 circulations over lifetime in collection
- Average title used 34 times more frequently online than circulated to a patron

**Example of a case where we hold both Print and Electronic versions of the same titles.**

**OSO was the first big product we purchased, continued purchasing the print – as we still do today. Seen as experiment, means of added access (full-text indexing to the print books) expectation that they would continue to want to actually use the printed volume.**

**Online used 17 times per year across the entire collection, print much less frequently (34 times less so).**

**I see no evidence of an increase in the print usage, although that was my original hypothesis.**

# Aggregator Titles

## NetLibrary

| Titles in Chicago Set | 6,172 |
|---|---|
| Use in 2008 | 52,367 |
| Used 100+ Times | 79 |
| Used 10-99 Times | 1,073 |
| Not used | 1,611 |
| Average use per title | 8.48 #1=Freakonomics (779) |

## Dual holdings – n=3,278

| No use of print or online | 473 |
|---|---|
| At least one use in both | 1,198 |
| Used in print but not online | 356 |
| Used online but not in print | 1,260 |

Example of purchasing of big bundles of content – only a handful of these titles were individually selected.

Not a huge area for us – although this is not true for many libraries, but still fairly impressive usage.

Average = 8.5 uses per year – less that OSO, but still greater than print resources.

As with most products – a trend we'll see – a small number of the titles account for most of the use. In this case, a single title "Freakonomics" had 779 uses.

Also typical, of the books we also hold in print, in about 15 percent of the cases where we hold BOTH the print and the online versions NEITHER was used in 2008.

Beyond those no-use titles, the usage of print and online is all over the place, with no real trend – some were used both place, and some only one place or the other.

STM

Springer E-Books
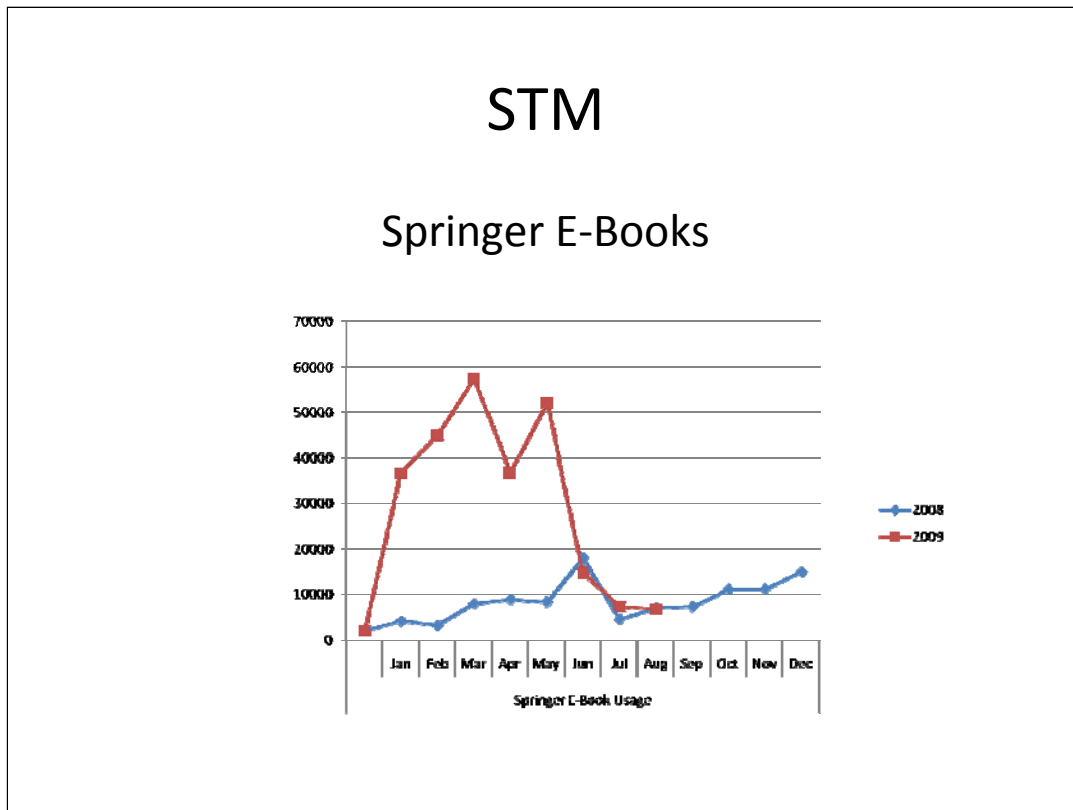
This next stide is an example of a case where we HAVE eliminated print versions.

This is the first time we've completely cancelled print holdings and are relying on online access only.

Large jump begins when we loaded MARC records into our catalog. Patrons are used to finding books in the catalog. DISCOVERABILITY We also know that MARC records aren't the only factor and we'll seeing significant use through traditional indexes, our link resolver and discovery in Google.

Usage is very large and encouraging.

Zero patron complaints, but some level of initial patron confusion since MARC records were delayed – resulting in patron requests to purchase titles we actually held in the electronic version.

Very happy with the Springer purchase on demand model, and see this as a crutch that made us more comfortable in our decision to get rid of the print.

# Technical Books

## Safari Books Online

- 1,143 titles in e-collection
- 26,201 uses in 2008
- 23 uses per title
- Top 2 = 10% of use (Learning PERL 2 editions)
- Top 25 = 50% of use

**Example of a wildly successful collection of technical e-books including many manuals which lend themselves to online use – often missing from print collection or charged to another patron, print use requires stopping what you're doing to walk to the library.**

**Publishers might want to think about content they have that would fall into this same "heavy use, immediate need" category and find a way to deliver that independently of their other content.**

**Again, we see the same pattern, where a small number of the books generate most of the usage. This is true everywhere, and directly parallels our experience with e-serials, where I have access to about 75,000 titles, but only a couple hundred of those generate 50% of my total usage.**

**This is something we should talk about more.**

# Major Reference Works

## Credo Reference, Oxford Reference Online, Gale

| Product | Titles Purchased | Uses in 2008 | Average per title |
|---|---|---|---|
| Credo Reference | 233 | 3,156 | 13.5 |
| Oxford Reference Online | 215 | 9,413 | 43.8 |
| Gale Reference | 131 | 32,550 | 248 |
| OED (Hits) OED(Full content units) | 1 | 5,151,691 256,458 | |

Have no knowledge of print use here since reference materials don't typically circulate. Only the Gale titles are fully cataloged in our collection – cataloging seems to drive use to a certain extent. These are all sets of individual titles, and there is no compelling reason for patron to search randomly through these collections. Highlights importance of access.

# Retrospective Collections

| Product | Uses in 2008 |
|---|---|
| EEBO (Document views) | 19,749 |
| ECCO (Section requests) | 18,437 |
| MOME (Section requests) | 8,337 |
| MOML (Section requests) | 272 |

Retrospective content plays an important role in our collections, and these four collections are all largely monographic in nature. Will we be asking for retro monographs the same way we were for retro serials?

# Textbooks
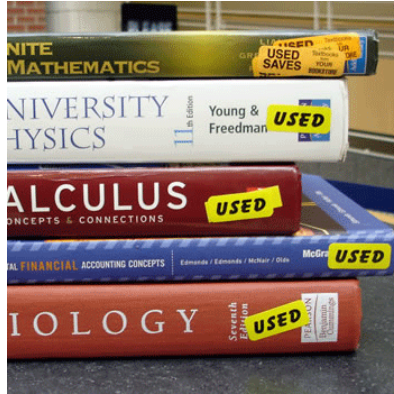
Not big for us, but interesting things are happening, experiments in online delivery. Some are opening say the textbooks need to be "open".

# Patron Initiated Purchasing

- Being widely discussed in library circles. Conflict between:
  - Building collections
  - Spending money "wisely" as need arises
- Assumes availability of the long tail titles
- Some libraries spending up to 50% of their monographic money on Patron Initiated Purchasing
- Print on Demand models

**Bullet 1 – Being discussed with a religious fervor that is coming close to equaling the great Open Access Debate, with people falling into two camps….. The model that is developing is one where libraries pre-populate their catalog with aggregator MARC records and they "buy" the book when someone clicks on the access button. That simple, with, I suspect much paddling below the surface.**

**Bullet 2 – The long tail extends to truly retrospective content – things like Gale's 18th Century Online, which is largely monographs. These older materials receive heavy use, in my library in the range of hundreds of thousands of uses for each product each year.**

**Bullet 3 - …and some are saying – perhaps provacatively – that in 5 years that will be 100%. Let me repeat…..**

**Bullet 4 – We're seeing several interesting POD models develop, including the Springer program, Espresso, Amazon, Google, etc. The concept of "no book will ever by out of print" is a tantalizing thought – one that will have a huge impact on front-list sales.**

# Buying a print book vs. buying an e-book

- One by one – or brought in via approval plans
  - Shelf ready in many cases
  - Approval plans finely honed for effortless purchasing
- Depend on notification – of both editions at the same time
- Different options for buying and distribution streams – one-by-one, aggregators, etc.

**All of these interesting things – like Patron Initiated Purchasing –assume a new efficient purchase model that isn't yet in place.**

**Bullet 1 – our print models emphasis efficient processing and easy flow of information….**

**Bullet 2 – this only works in the presence of notification upon publication, and really needs to be for all editions at the same time.**

**Bullet 3 – the many different options for purchase are leading to confusion in the industry at the moment.**

# Buying a print book vs. buying an e-book

- Subscription or purchase
  - Package or Title-by-Title
  - Rental? Rent-to-own?
  - Purchase on Demand – with preloaded records
  - Bundling with print – what if not purchased at the same time?
  - Consortial "sharing" of single copies
- Interlibrary Loan

**Some of those models include……**

**ILL – Purchase on demand (a sub-category of PIP, but this has been happening for awhile already), also Policy for e-books. We carefully write in the right to ILL, but no one really knows HOW to ILL an e-book!**

Content must be available where our patrons want to use it.

**How many of these do you recognize?**

**We really need portability and universality**

**To steal a line from Ranganathan who's basic principal was "For every book its reader, for every reader his book" do we now have – "For every ebook its reader, for every reader her ebook?"**

**Academic libraries were thrilled to get rid of multiple product-specific PC based reader software – now its back, but with different devices. A majority of our use is the delivery of content to a standard PC-based platform. While we have no problem with content being delivered via many platorms, we don't' want to find ourselves in a situation where content is linked to specific readers. This is unsupportable in the long-term.**

# Discoverability is Key

- Catalog (MARC) Records
- Indexing wherever practical
- Ubiquitous metadata
- Library initiatives to provide links
  - 11% of accesses via SFX link resolver (down from 23% in 2004)
- The Google effect and new initiatives
  - Serial Solutions Summon

**Our major goal as consumers is usage.**

**Last year my library delivered ~5M journal articles online = 416 per patron**

**E-books typically have a much lower overall usage, which mirrors the print world, but we've seen earlier that online use of monographs can be significant and seems to be greater than that of the print.**

**MARC records for e-books are a requirement, but we're seeing usage coming from a variety of places.**

**We all know abou t the "Google effect". Summon is another interesting new partnership – (With apologies to any Proquest people in the room for this vast oversimplification)  - in a nutshell, publisher metadata, along with our catalog records and other data, is all placed into a big Summon hopper and custom delivered to my patrons based on my holdings profile. This is exactly what many libraries have been asking for – a Google-like interface, with the content customized to what we have access to.**

**All these examples say to me that metadata must be free, fluid, maleable, linkable, portable, and sharable.**

**The stuff we make discoverable is the stuff our patrons use, and money spent to increase discoverability is money well spent on both sides.**

# ROI

- Increased emphasis being placed on analysis
- Long chain from initial use to documented ROI
- Analysis tools not yet up to the task
- Initial efforts
  - Paula Kaufman (Illinois) - $$ spent vs. grant income
  - Carol Tenopir – recent grant to study ROI

**Although this will seem like a joke to many of you, libraries are finally beginning to talk about ROI – partially under pressure from our parent organizations.**

**The days of simply buying everything and knowing that some of it is useful to some of our people are past.**

**Problems with the length of the chain, and poor analysis tools.**

**Work by Paula Kaufman and Carol Tenopir. Little experiments at Chicago, finding some interesting clusters of impact working with ISI and Scopus. Need to work together to get more information**.

# E-stuff isn't any cheaper

- Neither to purchase or to maintain.
- It does provide difference purchasing opportunities
- Serial pricing has settled down, e-books not so much

# E-stuff is better

- Serial Conversion [almost] complete
- E-books rapidly gaining acceptance
- Both offer the opportunity to explore new models of purchase and access
- The day of monolithic library collections is over



**Read list.**

**In summary, my question for you as publishers is how these sea changes will effect your business models.**

**This is an exciting time for all of us. It has been a great pleasure to share my thoughts and experiences with you. I hope this has raised a few questions, and I'm looking forward to continuing these conversations. I'll be around for the rest of the day as well as for most of the book fair, please to feel free to grab me.**

# Thank you

James Mouw

Director of Technical and Electronic Services

The University of Chicago Library

mouw@uchicago.edu

**Thank you very much for your attention, and now over to our next speaker.**