

Welcome to Stanford



Photo credit: e wood on fotopedia

Welcome to Stanford, or as we call it this time of year--Paradise.

From what I understand, all credit for this excellent weather goes to Kathy Long.

Thanks also to Kathy for asking me to speak this morning.

I'm thrilled to be here and grateful for the opportunity to talk to you all today.



Let me start by telling you a bit about who I am as a way of giving some personal context to what I am about to say.

First, let me tell you who I am digitally – so that anyone blogging or tweeting can properly credit or discredit whatever I say.

My blog is Feral Librarian, but as you can see, it is not anonymous.

I always have to explain the 4duke part of my twitter handle, and it is not only so that any crazy things I tweet get blamed on Duke not Stanford.



Photo credits:
Flickr user ebright
Bluedog423 at the wikipedia project

I went to Duke for my undergraduate degree and am pretty sure I signed up for Twitter during basketball season, because the truth is that during my 4 years at Duke, I spent way more time here, than at the library. So here is my first warning – be cautious about trying to predict propensity for librarianship based on library usage as an undergraduate. Some of us are late bloomers, and make our way to the profession through unpredictable paths.

I was in Army ROTC at Duke, so I went straight into the US Army as a young Lieutenant right out of college. I was stationed in Germany for 4 years in the late 80's and doing my part to fight and win the cold war.



Photo Credit: Army Sociology on Facebook

After that the Army sent me get a Masters degree and then here to teach Sociology and leadership to cadets at the USMA in West Point. I fell in love with teaching, scholarship, and Higher Education generally while there, so I resigned my commission after my 3 year tour, and came to Stanford to pursue a PhD in Sociology.



A bit of a contrast, but not much.

While in the PhD program here at Stanford, I took a part-time job doing statistical software consulting in the library ... when I finished my PhD, my library colleagues convinced me to apply for the Social Sciences Curator position that had been open at that point for 2 years. I thought it sounded fun, and I wasn't ready to leave paradise, so I applied, and got the job. That was 10 years ago, and I'm still not ready to leave paradise. You may have noticed that there is no mention of Library School or an MLS in my history – In that way, I am not a real librarian. So now you know why the blog is Feral Librarian.

My current job with the Stanford libraries is the AUL for Public Services – I am responsible for all the Social Sciences & Humanities libraries and librarians and for our Special Collections and University Archives. Let me tell you – that's a lot of humanities responsibilities for someone with degrees only in Sociology ... and from one of the most quantitatively and statistically rigorous sociology programs in the nation at that. What I have learned from my amazing colleagues about the humanities, about humanities research, and about library support for the humanities has very much informed my evolving perspective on libraries.

ROI: My concerns

- Short-term vs. Long-term
- Quantitative vs. Qualitative
- Accounting for rare events

So, what is my perspective on libraries, and specifically on applying ROI techniques and methods to assessing the value of libraries. As you might gather from the title of my presentation, I'm skeptical. I'm particularly skeptical about applying ROI to the work of large US research libraries like Stanford, but I'll be interested to hear if anything I say resonates with those from other types of libraries.

For me, an ROI framework is dangerous for Academic Libraries for 3 big reasons:

ROI tends to focus on the short-term & quantitative; and real impact of academic libraries tends to be long-term & qualitative.

An ROI framework doesn't account very well for "rare events". And I think Academic Libraries are about, at least in part, facilitating rare events.



Photo credit: flickr user [epSos.de](https://www.flickr.com/photos/epSos.de/)

Let's start with the short-term v long-term tension. When we talk about ROI for higher education, especially for research universities, we really aren't talking about economic returns. Research universities and Academic Libraries in the US are non-profits, so strict financial returns are not really our thing. To understand and assess the value academic libraries bring to universities, I think you have to look at the mission of the university— which is not about making money.

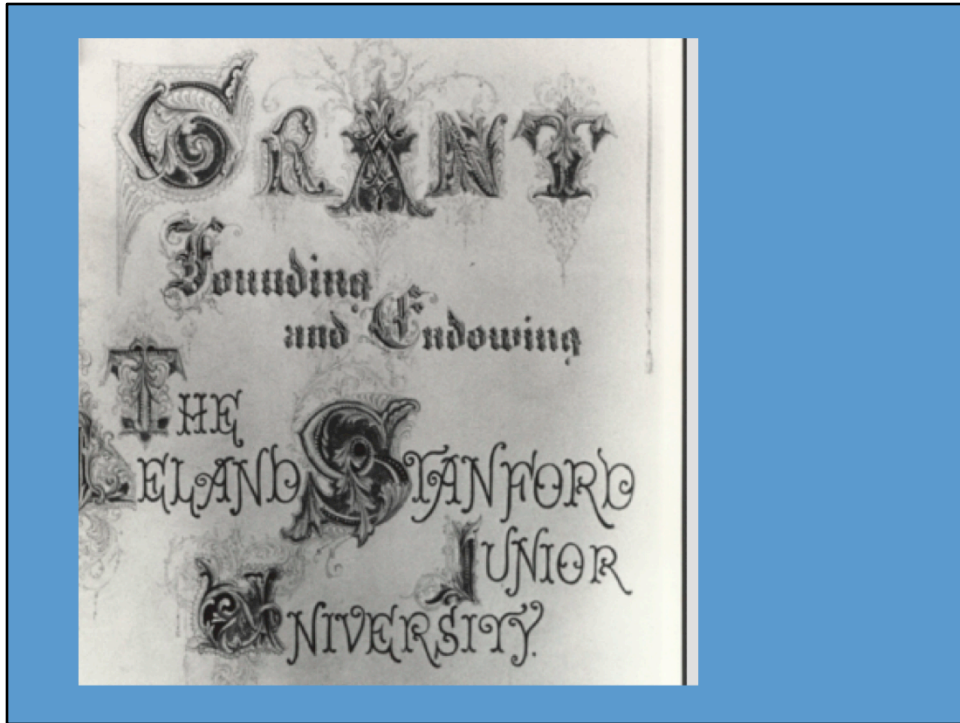


Photo credit: Stanford Historical Photograph Collection, <http://collections.stanford.edu/images/bin/zpr?cid=00007190-16470&fn=1>

So let's start with a look at Stanford's mission – with the caveat that Stanford doesn't actually have a current, officially labeled Mission Statement document. So we have to do a little archival research and look at our Founding Grant -which is preserved, discoverable, and accessible in print and digital form thru the Stanford University Libraries!

*“Its object, to qualify its students for **personal success, and direct usefulness in life;**
And its purposes, to **promote the public welfare** by exercising an influence in behalf of humanity and civilization, teaching the blessings of liberty regulated by law, and **inculcating love and reverence for the great principles of government** as derived from the inalienable rights of man to life, liberty, and the pursuit of happiness.”*

So the key original objectives for Stanford are to prepare students to be successful and useful, to promote the public welfare, and to inculcate love and reverence. Well that should be fairly straightforward to measure, right? We just need to see if the investments we make in library services are contributing to students “personal success and usefulness” and whether our collection development work and our digital library infrastructure development “promote public welfare”. That sounds easy enough, right?

I’m clearly being more than a bit facetious, but the truth is that deep down I do think that the investments we are making – in research services to students and scholars, in building deep collections, and in development next generation digital library tools and infrastructure do advance those original aims of the university, just not necessarily in tidy yearly increments that can be measured and reported as metrics used to gauge Return on our Investments.

*“Our goal for The Stanford Challenge is nothing short of building a university for the 21st century and beyond: A university that will **better serve the world** through the quality, impact, and vision of its research, and through the **new generation of leaders** it will produce.”*

~President John Hennessey

But moving on. Let's see if we can find something more recent, to see if the current university administration here at Stanford give us some more measurable goals to work toward.

This might work -- In kicking off the recently concluded Stanford Challenge—a huge development project that brought in loads of money--Pres. Hennessey articulated some goals:

“Our goal for The Stanford Challenge is nothing short of building a university for the 21st century and beyond: A university that will better serve the world through the quality, impact, and vision of its research, and through the new generation of leaders it will produce.”

We're still talking about some really long-term goals here. Basically, we want to produce great leaders, and solve world problems. Stanford's ultimate goals are not higher graduation rates, retention rates, or even higher employment rates for our graduates. So the returns on Stanford's investments, or even the return on any given students' investment in Stanford, can't be measured that way. Because those measures won't tell you much about how well we accomplish our actual goals – which are lofty and inspiring and long-term.

Library Mission

*** Collect**

*** Protect**

*** Allow use**

Academic Libraries, of course, exist to further the goals of their parent institutions. At Stanford Libraries, we support Stanford's goals and missions by doing what libraries do – we collect, describe, interpret, share and preserve information. As Archivist of the US David Ferriero recently described our work – we are in the business of “protecting, collecting, and allowing the use of information.”



Photo credit: Bighappyfunhouse.com

We do that in all the old traditional ways -- for example we still purchase about 130,000 physical books/year; and we logged over 628 thousand circulation transactions last year.



Photo credit: Chris Bourg

We also answer nearly 154 thousand reference questions and conduct 1000 workshops every year. And those numbers are not going down. At least at Stanford, the death of reference is greatly exaggerated.

On the surface, I think those sound like pretty impressive numbers, but when I really think about it; I am hard-pressed to articulate exactly how lots of reference questions and lots of check-outs actually contribute to a better world or to developing great leaders.

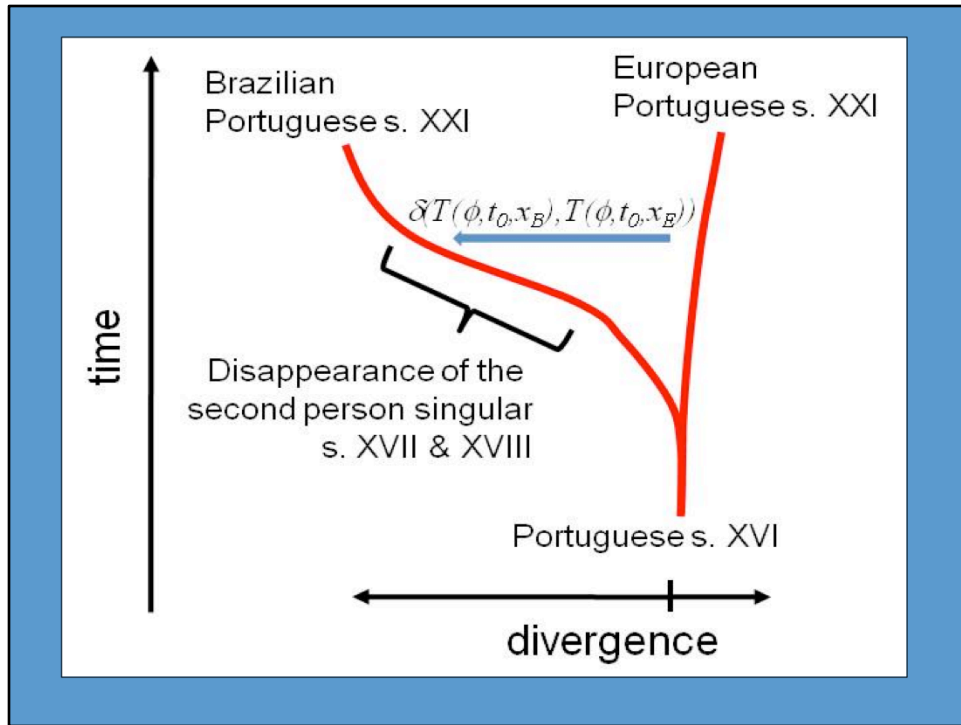
We also protect, collect and provide access to info in many new ways; as information production, discovery, use, re-use and consumption is happening in all kinds of new and innovative ways. Ways that our traditional measures of usage may not capture very well.



But of course, Stanford is more than a traditional library – we are proud of our reputation as leaders in digital library innovations. One, now sort of old, example of Stanford’s Libraries’ innovative spirit is the Google Books project. In 2005, we took a leap of faith as one of the original 5 libraries, agreeing to let Google digitize our collections. We did this in the hopes that getting the words inside our books indexed and therefore searchable would enhance discovery and would open up the treasures of our collection to a broader audience. Since the project began in 2005, the digitization project has grown to include nearly 30 libraries and over 10,000 publishers and authors from over 100 countries.

While we are disappointed in the lack of a settlement agreement to that would have allowed greater access to the full text of those books, we are still pleased with the impact of Google Books on discoverability and even use of collections.

The explosion of text-mining research that Google Books and other mass digitization projects enabled absolutely depended on Stanford and other libraries collecting and retaining a bunch of books -- many of which never or rarely circulated. New research questions can be asked now because the words in the books are no longer just words – not that there’s anything wrong with just words– but those words have now become data. And the sheer size and the near comprehensiveness of the corpus of digitized texts now available is only possible because libraries collected and preserved lots and lots of books over many years.



Graph credit: Cuauh Garcia

For example, there is a grad student here at Stanford using both Google Books and HathiTrust files of Portuguese language publications in the PD, as a means to tracing the evolution of Brazilian Portuguese. I suspect many of the books that are now being used as data have rather dismal circulation histories. If libraries like Stanford had only collected and preserved books with immediate and measurable use, the ever growing corpus of digitized texts would be even more skewed and biased than it already is.

In addition, I'm certain that the librarians who selected many of the titles that are now part of this scholar's data never anticipated this sort of use for their selections. Much of what we collect, we do so on a "what if" or "Just in case" basis; and we ought to be very cognizant of the fact that the objects we collect today (physical or digital) will almost certainly be used in ways we cannot yet anticipate.



Image credit: SPDO/Swinburne Astronomy

We are also making consider investments in collecting born digital materials – 2 areas of particular interest to me are Data Curation and Digital Forensics.

In terms of Data Curation--the amount of data that is being produced today is overwhelming. And I'm not just talking about the International Square Kilometer Array telescope due to be completed in 2016 that will generate approximately 1 exabyte of data per day, twice the amount that the global Internet produces in an average 24-hour span.

I'm also talking about massive amounts of social data and business data--which are increasingly one and the same. Just as there have always been more written words than we could possibly collect and preserve, there is now way more data than we can handle.

And while the Library of Congress announcement over a year ago that they would archive Twitter was met with great skepticism both in side and outside academia; we all know that social network data is invaluable to researchers in social sciences, marketing, and business.

My colleague Ron Nakao will talk tomorrow about what we are doing in the area of data curation and data management, so I'll only note that this is again an area where the ROI will be hard to predict – some data we can easily predict will be heavily used; other data ... we collect and preserve so that future scholars can make use of it in

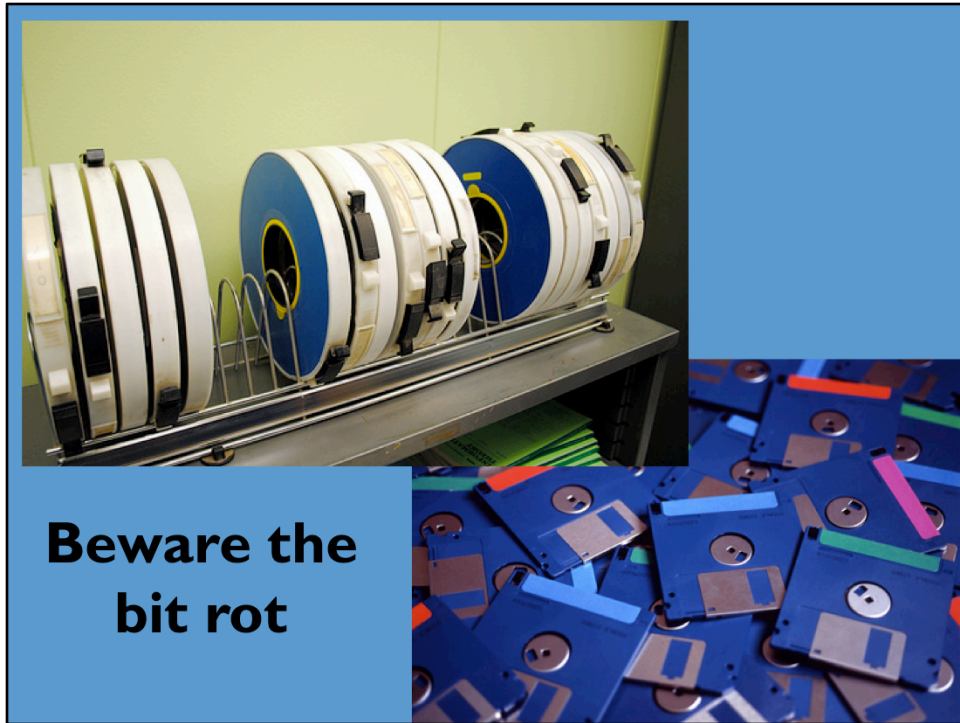


Image credits: flickr users FastLizaed4 & kickstock

As we continue to collect archival materials – based on our judgment about what and whose archives will have long term value to scholars and to society – we are increasingly collecting items that are born-digital. The Stephen Jay Gould archive is a great example – contains 60 5 1/4" diskettes, 81 3 1/2" diskettes, 3 computer tapes from 1987, 1988, 1994. The diskettes contain bibliographic databases and working drafts of many of the eminent author's publications. The data tapes appear to contain datasets used in his evolutionary biology research.

A key difference in investment in digital archiving and forensics vs paper archiving – with fairly minimal intervention, we can take a box of letters or paper manuscripts and put them in appropriate storage conditions and trust that when we get around to processing them, they will still be usable. This is sometimes referred to as preservation by benign neglect. That doesn't work with digital archive. The Gould discs and data tapes are already at risk -- we have to deal with them quickly, before the data deteriorates to an unusable state. So the investment in extracting, preserving, and reformatting born digital materials is often considerable.

And for much of what we collect and preserve, that investment represents a leap of faith. We are making our best guesses (as librarians and archivists

“scholars have uses for archives that archivists cannot anticipate.”

“The coolest thing to do with your data will be thought of by someone else.”

In collection development and in preservation, we have to focus on the long-term. I think we would be wise to keep in mind – and to remind our administrations – that we collect and preserve for today’s scholars as well as for future generations. This is true for traditional book collecting, and is true for new collection areas, like data and multimedia objects. Digital Humanities scholar, Dan Cohen, Director of the Center for History & New Media at George Mason Univ. says: *“scholars have uses for archives that archivists cannot anticipate.”*

Rufus Pollock of the Open Knowledge Foundation says “The coolest thing to do with your data will be thought of by someone else”

OK – so any assessment of the ROI on the collections of a large academic library has to account for long-term impacts. Perhaps that is not particularly controversial to this audience, but I do hope you are all preaching that story to your university administrators, your faculty, your presidents and provosts.



That said, I have already bragged about the fact that we continue to answer over 150 thousand reference questions a year, and that that number is actually up from last year. I also often brag about how much traffic we see in the building – as evidenced by the mountain of bikes parked at our entrance during exam periods. But before you let me celebrate that figure – and believe me I want to and I have – let’s stop and think about what it really means? This is where I want to highlight the quantitative versus qualitative tension.

How do the number of reference questions asked and answered help us understand our impact? Is number of reference questions a good proxy for the effectiveness of our reference program at contributing to teaching, learning and research at Stanford? Probably a better measure than some, but still not a very direct measure. After all – one could argue that high number’s of reference questions just means our online tools are too confusing (they are), that our website is not particularly user-centered (it isn’t) or that it is hard to negotiate our stacks (it is).

But, I happen to know that we have been promoting reference in our instruction sessions, our tours, and our workshops – so maybe rising reference numbers are a decent way to measure how effective our outreach has been.



Instruction is an area where I think we are doing a pretty good job at assessing our effectiveness. Here at Stanford, we have partnered closely with the Freshman writing program for many, many years – providing a designated librarian and a library tour & workshop for every first year writing class.

When I first took responsibility for the instruction program, our only assessment of this rather significant investment of librarian time was a survey we asked the students to complete at the end of their workshop. Those surveys were nice – we usually got high marks, and most of the librarians could use the results to learn how to improve their presentation styles (lots of us talk too fast, apparently); but those surveys told us nothing about how the workshops, and the availability of an assigned librarian contributed to the goals of the 1st year writing program, or to our goals of developing students' enthusiasm for and skill in finding and using scholarly resources.




So, we added a survey at the end of each quarter, after students have submitted their research papers. And the results are quite encouraging:

99% of students use the library catalog; and although we don't have any comparative data – we don't withhold the workshop from a random sample of students in order to have a control group – I feel pretty confident that 99% is a much higher percentage than we would get w/o our workshops. Moreover, of all the tools students used for their paper, the Library catalog and the Library databases were rated most useful (ahead of Google and Wikipedia). Nearly 40% of students consulted a librarian about their research paper. My favorite part of our results is that students who consulted a librarian were more likely to use Library databases and the online Research Guides than students who didn't consult a librarian. In addition students who consulted a librarian rated the Library databases more useful than those who did not.

To my mind, this is pretty good data to show that our investment in instruction is paying off in terms of use of library resources and an appreciation of the value of library resources – including the librarians themselves.

The next logical step for us would be to conduct even longer term assessments – it would be great to know if the work we are doing with the freshman pays off throughout their four years at Stanford and beyond.



**Focus on short term
returns is dangerous and
misleading**

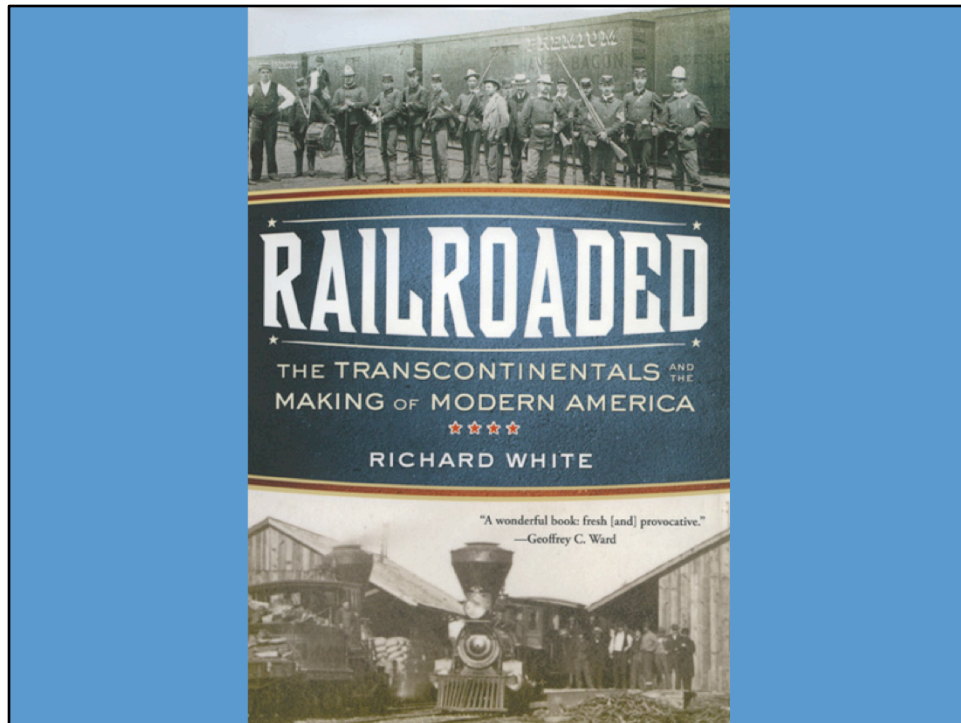
So, by now you can see that I'm not categorically opposed to library assessment or to the practice of calculating the returns on our investment. I am merely cautious about it, especially when an ROI approach leads us to focus on short-term outcomes that might be very far removed from the long-term goals we have of supporting research and learning in the service of developing educated citizens who will solve world problems. I think we need to really think carefully about the data that we collect and the metrics that we use, lest we start to mistake reference traffic or circulation statistics as the real goals.

In addition to a willingness to patiently focus on long-term returns when assessing value of academic libraries investments, I would argue that we also need to recognize the qualitative rather than merely the quantitative nature of our contributions. To my mind, circulation is one of the most over-rated quantitative measures that we have.



Photo credits: amazon.com; Kenneth Chan

When we look at Stanford Libraries' circulation numbers, the Lord of the Rings DVDs would seem to yield the highest return on our collection investments – since it is our most heavily circulated item in the last 5 years. Now, as perhaps the only American librarian who has actually never seen nor read Lord of The Rings; I feel that I must pause here and note that I am not saying we shouldn't have the Lord of The Rings or that we shouldn't be quite happy that it circulates. I'm sure they are fine movies, with considerable academic value. But I do have an issue with using circulation as a key measure of value, if for no other reason than it would lead us to overvalue Lord of The Rings and undervalue collections like our historical newspapers. After all, Lord of The Rings is our most heavily circulated item, and the microtext collection that contain the text and images from 100s of years of historic US newspapers are much less frequently used.



But if we care about actual impact on research, we might want to look at historian Richard White's recent book Railroaded, which provides a new and controversial vision of the the so-called Gilded Age in the US, and the impact of the Trancontinental railroads on the making of modern America. White's book reveals the chaos, dysfunction & failures of the early American railroad industry, and many reviewers note that there are lessons to be learned from White's research that are relevant to our current periodic of economic turmoil. White relied heavily on archival materials rarely used by others, and on dusty reels on microfilm that he may well have been the first to pull out of the file cabinets.

"the legendary Margaret Kimball helped me go through the holdings and find what I needed. Jim Kent, who runs the media and microfilm room at Green Library, and his staff helped me in ways probably best kept between us..."

The centrality of the archives to White's research is acknowledged, quite literally, in the acknowledgements section.

In fact, in addition to acknowledging by name the University archivist and the head of our Media MicroText center, White notes that "a great part of the pleasure of writing this book has been the time it allowed me to spend in the archives. The paradox of archives is that there, among the relics of the dead, the past seems most vital and alive."

I use this example because the key sources for Richard's research were almost exclusively low-use materials – materials that had yet to show much return on our investment in them.

I like this example for at least 2 other reasons:

1—it highlights again the long term nature of library work – as we all know, books take a long time to write. In this case, the librarians Richard acknowledged had both long since retired before the book's publication.

2—I think acknowledgments of libraries, archives, librarians, and archivists in published materials constitute a direct and real measure of our impact on scholarship. I don't know of any academic library that measures and tracks acknowledgements systematically; and I'm not even sure how it might be done. But I think it is an idea with considerable potential.

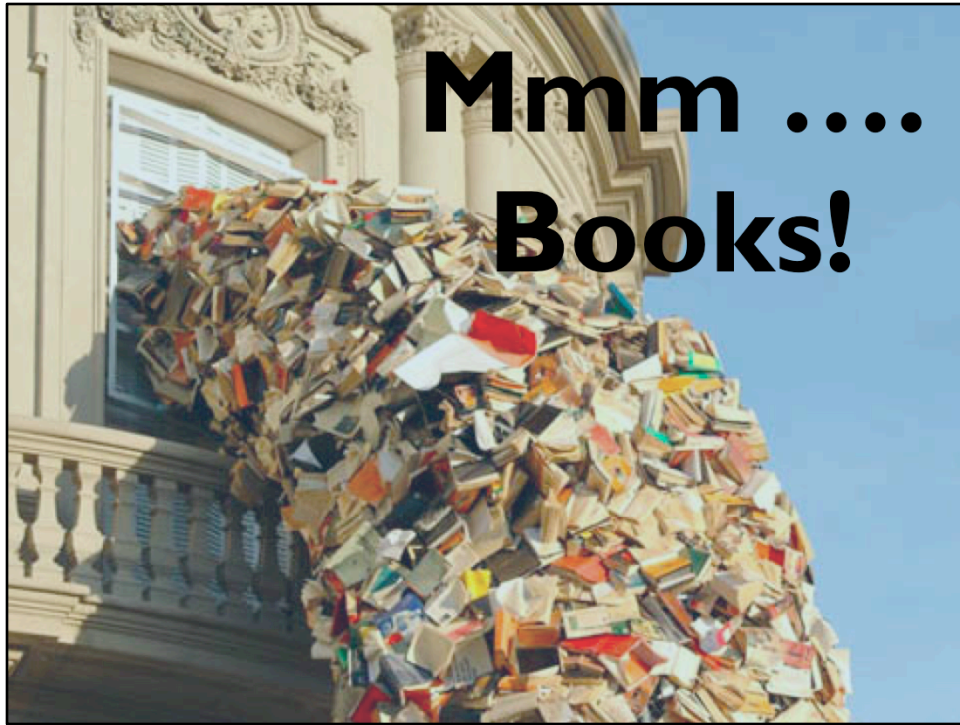


Photo credit: **Biografías** (2003), by Alicia Martín

Let me talk again about books – about the value of simply having a large collection of books, of written words—even if some, maybe many of them, don’t circulate. Students and scholars benefit from being surrounded by lots of books. Remember that picture of all the bikes? Students still rank the library as their favorite place to study – and we don’t even let them bring coffee in. They like being surrounded by book – even if they don’t read them.

My favorite description of the value of books comes from writer, editor, book reviewer Kristy Logan, who wrote recently about the impact of the 800 unread books on her shelves:

“Sometimes I hold these books in my hands and imagine what I will learn from them.

These books have affected my writing, and I haven’t even read

them. Maybe we can learn as much from our expectations of a story as we can from the actual words on the page.”

Books affect my writing, even if I haven’t read them.

And of of course we all know scholars who download articles they will only skim, who pulls books off the shelf, only to put them back (usually in the wrong location) after a cursory look) and even who request via Inter-Library Borrowing lots of book they don’t actually read and that never make it into the bibliography of the book they publish. But all of that not-reading is still part of the research process none the less.



And if these qualitative testaments don't convince you, there is data! Results of a cross-national study of family scholarly culture and children's educational attainment, published in 2010 found that:

"Children growing up in homes with many books get 3 years more schooling than children from bookless homes, independent of their parents' education, occupation, and class."

My point is this – there is value to libraries and to collections that are no less real and no less impactful for the fact that they can't really be measured. That is a hard truth for a quantitatively trained sociologist to admit, but I have come to believe it. So, yes, we should practice continual assessment and we should gather as much evidence as we can that shows the impact of our collections and services on the goals of our institutions – but we should do so with an eye towards the long-term with a commitment to valuing the qualitative as well as the quantitative impact of our efforts.

What about *Serendipity?*

So now let me turn to my real, real concern with ROI –which is even long-term qualitative assessments of impact will always miss one of the most important functions of an academic library – which is to facilitate the rare event, unexpected, dare I say random event. Yes, I know, how novel—a librarian talking about serendipity. But remember, I'm not a real librarian, and my belief in serendipity has developed slowly and skeptically – but I am a convert. And perhaps I preach w/ the over-zealousness of the recently converted. But I have come to believe that it is absolutely the responsibility of libraries to encourage, support and in all ways make possible the unanticipated discoveries that lead to new knowledge, new ways of thinking and new contributions to the cultural and scholarly record. In fact, I think providing the context in which new, unanticipated, unique discoveries, thoughts, connections, and inspirations are sparked may be the most important value-added contribution that libraries make. Allow me to share a couple of fairly recent examples of the kind of serendipity made possible by the careful work of libraries.

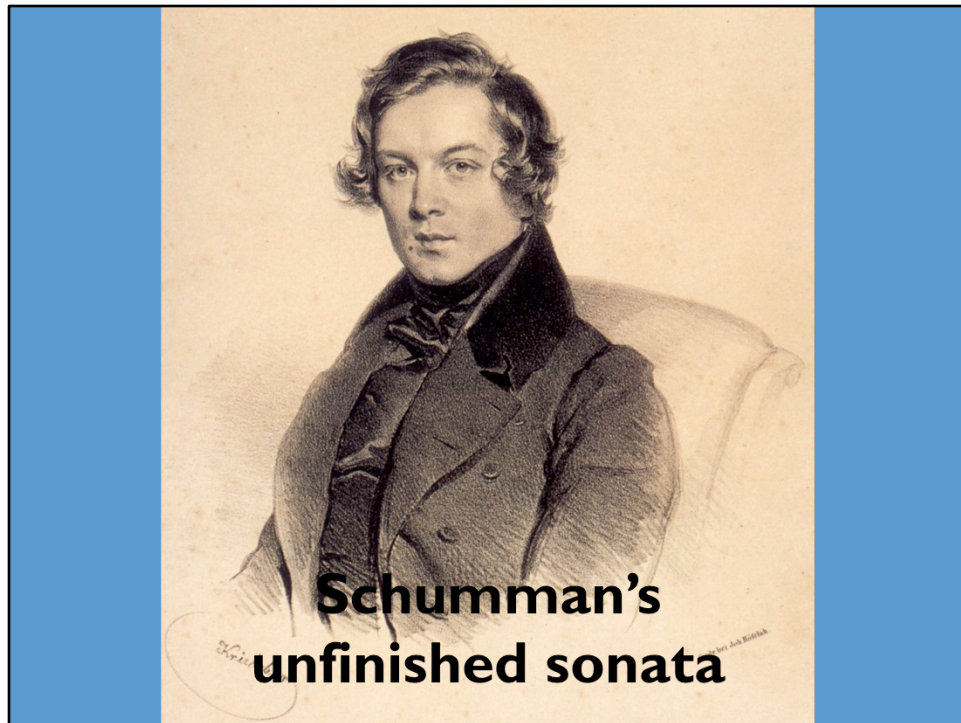


Image credit: Frederickmoyer.com

In 1989, in honor of Condoleeza Rice, Walter Hewlett gifted to Stanford an autographed fragment of a musical score from German composer Robert Schumann. Some 20 years later, Frederick Moyer, a concert pianist, and his uncle, Paul Green, an engineer; tracked the score down in the Stanford Libraries, requested and received via InterLibrary Borrowing a digital scan of the score, transcribed them, and entered the score into clean computerized format, and developed software that allows the listener to hear Robert Schumann's hitherto unfinished 4th sonata, as played by Moyer, while following along with the handwritten original score.



Zora Neal Hurston's lost stories

Image credit: U.S. Library of Congress

[\[1\], Reproduction number LC-USZ62-62394 \(b&w film copy neg.\). Card number2004672085.](#)

And here's another example:

In 2010, 2 Harvard professors "discovered", by scouring newspaper microfilm in the basement of the Harvard library, 2 new short stories by famed Harlem Renaissance author Zora Neale Hurston. These 2 stories were never listed in any of the published collected works of Hurston and had not yet been studied or analyzed by other scholars.

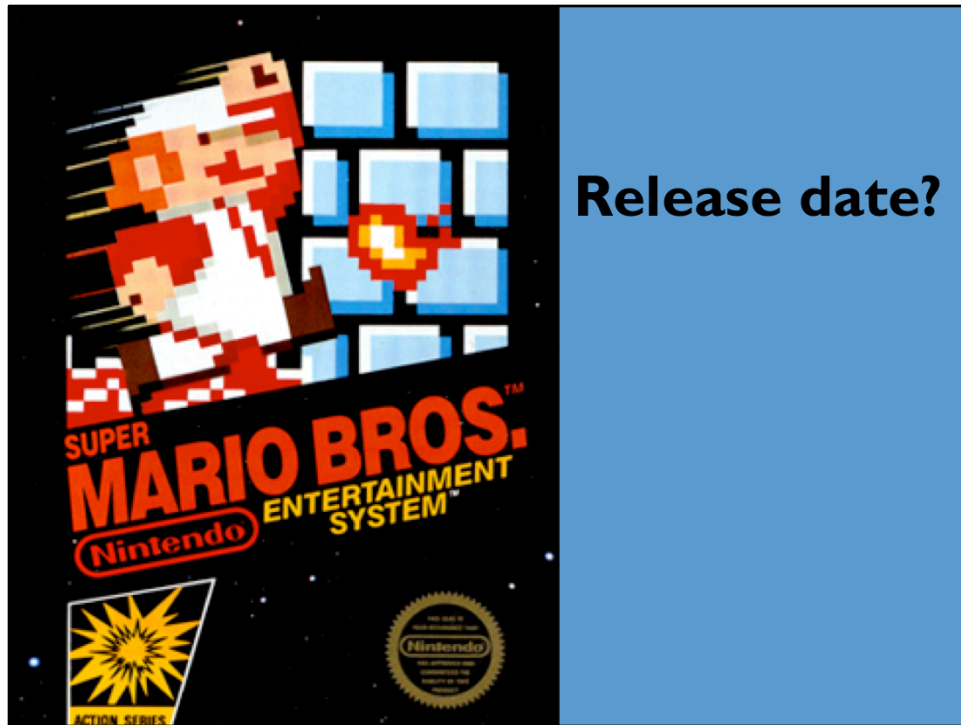


Image credit: Screenshot showing the box art of *Super Mario Bros.* for the NES from Wikipedia

And here's an example of the failure of the archives (archives writ large, not any particular archive). One of the great unknowns in video game history is the U.S. release date for Super Mario Brothers. We take video game history rather seriously here at Stanford, in fact we hold one of the largest historical collections of video games in the world. The lack of careful documentation and archiving of that documentation actually represents a fairly substantial gap in the history of video gaming, as Super Mario Brothers is one of the most successful, iconic and influential video games in the history of the industry. But it is hard to confidently trace its influence on the development of the industry when historians can't yet agree even on the year of its US release.

***the plural of anecdote
is not data ...***

...or is it?

I'm sure many of you could provide other powerful examples of the same sorts of serendipitous discoveries of and rare uses of materials hiding in libraries. Or of missing archives that hinder scholarly progress.

And yet -- in the back of my head, I hear a little voice reminding me that "the plural of anecdote is not data"

Except that maybe it is ...

Because let's all remember that the fact that these accounts are merely anecdotal does not render them any less true. These stories, and countless others, represent real contributions to scholarship and to our understanding and appreciation of the world.



***“...chance favors
the prepared mind”***

Image in the public domain

Serendipity is by definition an unanticipated, but happy, chance occurrence. Perhaps the most famous quote about serendipity comes from French scientist Louis Pasteur, who said “in the field of observation, chance favors the prepared mind.” I always like to add what I call the librarian clause to this great quote: “Chance also favors those with access to a great library.”

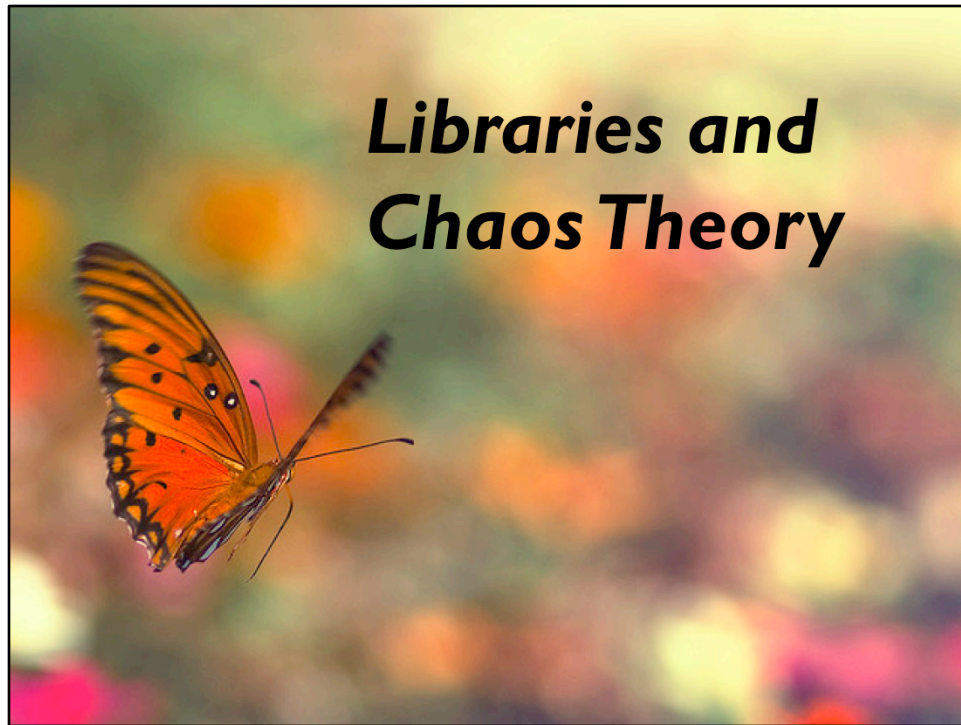


Photo credit: flickr user Joel Olives

As long as we are talking about chance, allow me to throw in a hopefully not so gratuitous reference to the next talk by Jennifer Aaker (author of the Dragonfly Effect). It is increasingly clear to me that what scholars and students find and do and learn and discover in libraries and through libraries is often one small link in a chain of events that may ultimately lead to ground-breaking research, to transformative art and music and other creative expressions, or to the growth of a student into a great leader, perhaps even one who has developed a lifelong “*love and reverence for the great principles*” of democracy. In other words, encounters in libraries-- with books, with digital objects, with librarians, with other students and scholars – can be seen as a sort of scholarly flapping of the butterfly wings – perhaps small acts or brief encounters, but with far-reaching and unpredictable effects. And if we want to understand and capture the true returns on our investments in libraries, we have to be willing to account for those small, chance encounters.

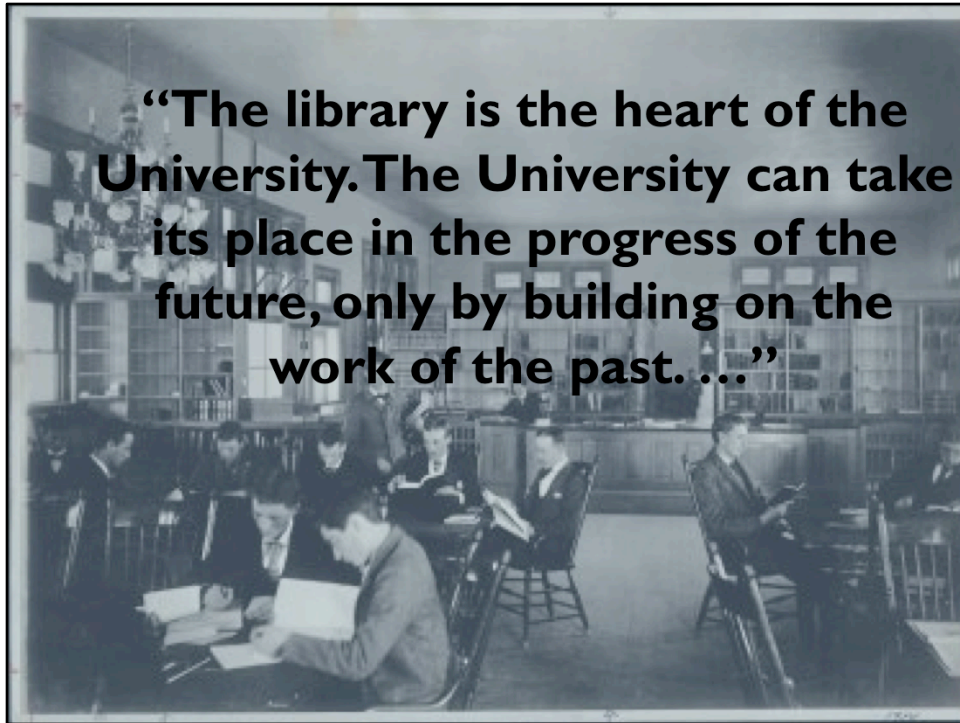


Photo credit: Stanford Historical Photograph Collection

A final thought or two --from Stanford's 1st President, David Starr Jordan who understood and articulated well the real and the symbolic value of a library.



Perhaps I have presented an overly romantic, even mystical portrait of academic libraries – and at a time when libraries and higher education are under the gun to get practical. But what I am suggesting is that if we don't defend the hard to define and even harder to measure qualitative importance of libraries, who will?

And, I suspect that many of you probably agree with me, at least in principle, that universities ought to have great libraries, with expert staff and awesome collections and a range of services in support of teaching and research. But of course, we all face constraints in the forms of budgets, space, and competing priorities.

So, yes, by all means find good ways to measure our contributions to the aims of higher education. But also, please, take opportunities to evangelize on behalf of the immeasurable impact of libraries – make sure your administration knows that there is value in books that aren't read, in data that hasn't been used yet, in archives yet to be discovered, and in the mere fact of great libraries.