Against the Grain

2015

Pelikan's Antidisambiguation--"Self-Preservation and the Cloud"

Michael P. Pelikan
Penn State, mpp10@psu.edu

perpetual access, multi-viewer, on the first view that exceeds five minutes for a price of $750.00. Then couple this profile with a separate profile for all other video that states: subscribe for one-year access, multi-viewer, on the third view that exceeds 30 seconds for a price of $99.00. And you can extend this logic to high-use e-textbooks versus esoteric scholarly manuscripts. Or high-use audio tracks for music appreciation courses versus obscure performances on the Clavichord. And on and on… Without such a model, much high-use, high-demand content is likely to not be available in DDA.

Multi-Media

Why should a librarian or library patron have to access DDA content via multiple platforms and vendors? eBooks, archival documents, audio tracks, video, musical scores, data sets, journal articles, etc., etc. are all content types amenable to measurement, use and sale. The aggregation and distribution of content by small and large companies alike is increasingly multi-media and, therefore, the platforms must eventually also be multi-media. Specialized collections, especially in areas like music and film studies, provide scholars and students with a mix of media types for study. Of course, the *"Modern French Film Studies Collection"* can be purchased, and the individual items within the collection can be purchased via single-title sales, but DDA via a single, multi-media platform allows usage to determine the purchase pathways and the student of modern French Film might well be the trigger of the purchase of a video, a film script, a reference monograph on the film, and a biography of the director; or not… depending on the level of interest and the purchase trigger parameters selected by the librarian. Massive aggregation of eBooks, through platforms like **ebrary**, delivered DDA, has allowed libraries to migrate toward a single eBook platform and evade, even partially, a state of "platform weariness." Imagine then a future state where a single mixed-media DDA platform supports access to all the media types central to a student or scholars search within and across disciplines and areas of study.

Data Analytics

Seven years ago, when I founded **Business Expert Press**, our eBook collection was made available exclusively through the **ebrary** platform. At that time **Counter** statistics provided little more than title-level and collection-level numbers of views and total pages viewed; neither we at **BEP** nor the libraries that purchased our collection had visibility into usage beyond these raw statistics. I suspect our internal team spent more time reviewing usage statistics (to assess the likelihood of a given library continuing to purchase our collection) than did our library customers. Usage data, to be truly valuable, must be more robust. Near the end of my time at **Pearson Education**, the company acquired **E-college** and I recall the standout feature of **E-college** that made it such a desirable acquisition target for **Pearson** was the robust "back office" data they provided university administrators on online course, program and instructor efficacy; measured both in student results and program profitability (i.e., enrollment rates and completion rates along with costs to support a given class).

The state of data analytics in support of the classroom and the learning enterprise has become increasingly individualized to the learner and has moved forward at a much faster pace than within the library content space. Fast-forward to 2015 and the current state-of-play in DDA. At present we are giving librarians little more than data about pages read, minutes viewed, total views, total users, etc. What if a librarian was able to distinguish between the views of faculty versus students? Ph.D. students versus undergraduates? The type of content (learning as opposed to scholarly reference) viewed by which departments at which time in the semester? Discipline profiles? Correlations between online programs, degrees, and content triggered for purchase? Location of view: classroom, dorm, in library, off campus? Perhaps some of these possibilities strike a chord with you, the reader, and perhaps some of this is improbable and unnecessary. But evolving usage and user data that offers ever deeper insight into the values and needs of the library patron is the indispensable corollary to the description above about customization of the DDA profile; a library that knows a good deal about how its content is used will make ever better decisions about how to trigger purchase in an improved future-state DDA environment.

It is my hope that this column sparks debate, inspires publishers and aggregators, and sets off a conversation about how far and how fast we move with DDA. 🦅

Pelikan's Antidisambiguation — "Self-Preservation and the Cloud"

Column Editor: **Michael P. Pelikan** (Penn State) <mpp10@psu.edu>

Despite this column's dedication to the notion of "Ambiguation" (if you can disambiguate something, you must be able to ambiguate it, right?), I don't usually go in for "ambiguated" titles. We're making an exception this month, to reflect my genuinely ambivalent feelings about that most ubiquitous, aggressive, and downright inescapable of recent ideas to have had its turn at sweeping the "Net-mosphere": the Cloud.

I can't remember a Net-based concept that has achieved such a degree of pervasiveness more quickly and completely than the Cloud. Certainly there are ideas of similar or comparable scope; social networking, for example. And there are clearly individual products that have achieved sweeping adoption in a very short time; Facebook, for example. But remember, in many ways, Facebook is simply an individual branded meta-service whose existence relies on the presence of the Cloud to make its connections, to keep it running, and ever-expanding. Same story with Twitter. Same story with Linked-In. These are entities that go beyond simply being something on a server. In most ways that matter, they are platforms, existing in and reaching out from a virtualized setting. They are way more than just Webpages.

For a useful exploration of the origins of the term in question, I refer you to the Wikipedia article entitled, "Cloud Computing." I'll just touch on a couple of points here. Wikipedia says use of a cloud symbol to represent the Internet goes back to 1994. The phrase "Cloud computing," however, received a primary socializing boost in 2006 with **Amazon's** introduction of "The Elastic Compute Cloud." Indeed, the term coined by **Amazon's** marketing arm for its net-based virtualized computing platform, "EC2," comes from an initialization of the words "Elastic Compute Cloud."

The Wikipedia article aptly describes the use of the word "cloud" "…in science to describe a large agglomeration of objects that visually appear from a distance as a cloud and describes any set of things whose details are not inspected further in a given context."

This turn of phrase, in fact, distinctly characterizes that which is in common among the vast majority of Cloud-based services we use today. It's your "stuff," but the details associated with the management, storage, and retrieval of your "stuff" are not apparent, nor need you apprehend any of that to make use of the service. It's in the Cloud. Don't trouble your pretty little head about it.

From a practical perspective, it has proven very easy, very natural, to become accustomed to having immediate access to a wide variety of my "stuff," regardless of what device I was using when I first wrote something, read something, took a picture, or listened to a piece of music. Web history? It follows me around — if there's a spot I visited in Chrome on my tablet this morning, I can find and reopen it this afternoon on my phone. Easy. It just works. And it reaches across devices, platforms, even applications — even modalities. Just as this column was going to bed, there came the announcement that **Amazon's** Echo would now be able to read to you, out loud, from books purchased through your Audible account. The prospect of "Alexa" reading to me is persuasive — it was one of the missing capabilities I noted early on with the Echo. I was pretty certain the obstacles lay more in the realm of licensing than in the details of technical implementation — remember the flap over whether the first Kindles would be able to read books you'd purchased, or, excuse me, I mean to say, you'd licensed? (Old perceptions die hard.)

In terms of keeping my "stuff" safe, there are few more seamless examples than **Amazon's** cloud-based infrastructure behind the Kindle. Need a preservation strategy for old articles — the ones you used to photocopy and hang on to? Now you can simply render them into a pdf and email it to your Kindle's email address. Not only will it appear on your Kindle, but when you buy a next Kindle (which you will, or at least, I will) you won't have to copy a thing onto it — your "stuff" is already in the Cloud, waiting for you to download it. Books, articles, music, photos, videos, all are safely enfolded in the Cloud. The

Cloud is your self-preservation strategy.

Well, that's one kind of self-preservation. But what about the other kind of self-preservation?

If you start digging into the Terms of Service of a few of the big operations you're signed up with, it'll make you want to run away, hide, and spend your time washing your hands over and over.

From **Google's** Terms of Service: "When you upload, submit, store, send, or receive content to or through our Services, you give **Google** (and those we work with) a worldwide license to use, host, store, reproduce, modify, create derivative works (such as those resulting from translations, adaptations or other changes we make so that your content works better with our Services), communicate, publish, publicly perform, publicly display, and distribute such content."

It goes on, "This license continues even if you stop using our Services…"

What's more, "Our automated systems analyze your content (including emails) to provide you personally relevant product features, such as customized search results, tailored advertising, and spam and malware detection. This analysis occurs as the content is sent, received, and when it is stored."

That's right, **Google's** bots read your email. If I mention in an email to my astronomy buddies that I'm thinking about getting a new Cat, I'll probably start seeing ads for kitty litter next to my news. Of course, I was referring to a catadioptric telescope, known as a "Cat" to its aficionados. So maybe what's so annoying about the system is how it is sometimes so transparently dumb. Or maybe it's that I know perfectly well that before long the bots will have figured out the context of my "Cats" and I'll start seeing ads for **Celestron** or **Meade Instruments**…

I don't mean to single out **Google** here — not by any means. These constructions are not substantially different from many you encounter once you lift the lid and peer into what's happening in the Cloud. The first observation of possible import here is the degree to which we permit ourselves to become numb over steady, if incremental, erosion to our boundaries. A person with whom I was discussing this recently acknowledged the incremental "numbing down" of our sensibilities, likening it to a lobster being lowered into the pot of boiling water. It's not to say that it doesn't hurt; only that the hurt changes, travels, and ultimately doesn't hurt so much after a while. It's also difficult to "unboil" the lobster.

So powerful are the preservation capabilities built into the Cloud, and so seemingly irreversible, that something mimicking an entire social movement has arisen around what's being called "the right to be forgotten." The question we face is this: is there a business case for the cloud services to relinquish back to their users the rights the users so readily relinquished to the service? Without a set of reasons to do so that equate to "sound business," why would the Cloud services voluntarily hand over what are, in fact, among their most valuable assets?

Usually, I think, the Cloud services can quiet the conversation down simply by rolling out their next phase of product enhancements, restoring to the Faustian trade its rightful irresistibility.

I have a friend who advocates trading grocery store loyalty cards with friends, acquaintances, and even perfect strangers, every few months. He points out, accurately, that you still get the discounts. But more importantly, he says, it messes with the behavior modeling the cards provide data for. Just when the store was about ready to conclude that people who buy paper products with the word "Nature" in the brand name will also pay extra for oven cleaner with the word "Organic" on the label, along comes a fresh set of data indicating that, no, people who will pay extra for "Organic" oven cleaner also really like snack products that are colored neon orange and whose labels imply the presence of cheese.

This is a noble idea, and possibly useful, at least until the grocery stores catch on and start requiring additional forms of identification from their patrons. We'll give it to them too — and in fact, we probably already do. Who wants to receive coupons for products we are unlikely to buy? Much better to get coupons for things you care about.

I'm reminded of the Gamekeeper from the original Jurassic Park who, observing the Velociraptors hurling themselves against the electric fences, looking for weaknesses, points out, "They remember…"

Nobody remembers as long as the Cloud… 🦐

@Brunning: People & Technology

At the Only Edge that Means Anything / How We Understand What We Do

by **Dennis Brunning** (Director, The Design School Library, Arizona State University) <dennis.brunning@gmail.com>

Life as We Live It Random

Internet kids claim and use random. It's a put down. **Luc Sante** in a *New York Times* Sunday Magazine essay feature brilliantly places random in its exhausted etymology. Like so many words in our hyperactive worldwide Web vocabulary, it has had different usage. According to the 5th edition of the *American Heritage Dictionary* (love it to descriptive death), it means without purpose, design, or method. **Sante** understands this well but expands into its current connotative meaning as "not one of us."

You are "rando" if you cross my field of vision and I don't care. Or you are rando if you intrude on my space and time without introduction. Photo-bombed by a stranger is true rando.

None of us, to us, are random. None of our friends or family are random unless we define them as such.

Sante takes the tour of random (which may to you now be quite rando). It began from the Old French "randir" which meant "to gallop." *OED* indicates then it became associated with impetuosity, great speed, force, or violence." By the 20th century **Bennett Cerf** and **Donald Klopfer**, enterprising publishers, sought to introduce a line of books they'd issue at random. Thus **Random House** was born.

Sante delights in the 163 random definitions in the *Urban Dictionary*. It has become a nonce word (itself rando). He notes that anyone can post a definition to *Urban Dictionary*, yet another random act of unnecessary niceness. It reads like a perfect Eluardian nonsense like when the appropriate response to a ninth grader's exclamation "Elephants are contagious." The rando response: "that's random."

This is insider jargon at its best. You observe and announce what the world in front of you means for you, and it means nothing. Those who get it, agree that it's random, or they nod knowingly. It's defining the other not as **Sartre's** "hell" but just not you and you are center. Others are arbitrary and you are not. You have arrived and no one else.

However this is not so random as **Sante** or your teenager might think. Years ago, in the pre-Internet law library at the **University of Illinois** where I was randomly taking a course in government documents (how rando is that?), my all but doctorate instructor commented to us while rolling her eyes, that jet flight was so random. It was time travel and between departure and destination, between arrival and departure, among passengers, friends, and strangers, it meant little.

Later, after I endured the course, got an A, I called all but doctorate up for a date and we ate pizza at Papa Dell's on Green. We looked at the pizza, like any other pizza, we laughed about the SuDoc system, and we briefly looked each other in the eye. I left her place the next morning really not thinking about random but remember it today and every day since as an entirely random experience.

"Ah, Dad, that's so rando!"

Reading

This summer it's all about robots, driverless cars, shadow work, and no bosses. We range in emotion from **Marc Andressan's** "Software is eating the world" and the robotized underemployed future to how the same technology, done right, can free us. It is about the closed and the open. How you feel after reading depends largely on where you lie on this closed open continuum.

Craig Lambert's *Shadow Work: the Unpaid, Unseen Jobs that Fill Your Day*, explains how we are making ends meet despite slow uptick in employment and what the ungainfully employed are doing. We are working harder, longer hours, more jobs, that fill in the spaces