Sarah Potvin, Digital Scholarship Librarian, Texas A&M University Libraries
Notes for presentation to Texas STEM Librarians Conference
July 17, 2014, College Station, Texas

Slides for the presentation have also been posted to the OAK Trust institutional
repository

Slide 4.
Metrics are, of course, measurements. Or, to pull from Merriam-Webster, a metric is
"a standard of measurement."

When pulling together a CV, going up for promotion and tenure, applying for grants,
etc., what is it that researchers are trying to capture about their work? And what is it
that colleagues, university administration, promotion & tenure committees, and
granting agencies are trying to determine about the work? What data about a
scholarly object or researcher makes a case for or against tenure? Or promotion?

What makes a case against other awards, or demotions?

How can you judge the quality of scholarship?

What about those attempting to put the work into context, say for inclusion in a
literature review or just for basic discovery when researching a topic? Outside of the
academy, what is it useful for clinicians, practitioners, policymakers, and other
interested parties to know about an item of research?

In a basic, general way, we use metrics all the time when navigating the world—I am
very swayed by rankings on TripAdvisor and Yelp, for example. What metrics
become important when we are assessing a book, an article, a GoogleDoc, a
slideshow, a dataset? And how do these metrics tie into discovery?

And how do they help up manage our own work?

The Oxford English Dictionary offers up a nontechnical definition of impact: "Now
commonly the effective action of one thing or person upon another; the effect of
such action; influence; impression." What do we mean when we claim that a
scholarly item had an impact, and how can we measure it? Too often, the
measurement precedes the definition. Or maybe impact isn't what is being claimed,
or aimed for. Some researchers or audiences may prize highly accurate research,
high quality publications. What would evidence for these claims look like, and how
can annotations, stories, and numerical evidence help make the case about an a
research product's importance, quality, influence, distribution?
Related use cases:
Reward
Making the case for (or against) tenure/promotion.
Making the case for other awards (or demotions).

Discovery
Getting this research to the people who need to find it.
Acquiring this research.
Contextualizing the research.

Slide 5.
Discussion of how these are measured, and what the advantages and drawbacks are / metrics and some of the controversies and drawbacks:
Author-level
The H Index looks at the output of a particular researcher and attempts to capture productivity and impact. It is based on the distribution of citations of publications. Only works for comparison within a given field. Calculated by Scopus, Web of Knowledge, Google Scholar, and other databases. Changes depending on how database accounts for citations.
The i10 index looks at the number of academic publications by a given author that have been cited at least ten times. Developed by Google Scholar.
Item-level:
A citation impact can literally just be the number of times a given item was cited. Grouped with this, you might also consider page views or item downloads, i.e, data on usage provided at the item level.
The last two are journal scores.
Eigenfactor is intended as measure of the importance of a journal in its field. Weighted ranking based on the number of incoming citations (weighted by the relative importance of the origin of the citations). Not averaged out by number of articles, so bigger journals have higher scores.

Show *Seed Magazine* video on Eigenfactor: http://www.eigenfactor.org/about.php
Take a couple of minutes for an exercise on evaluating the Eigenfactor:
Pay attention to what the claims are that are being made in this clip– what are they including when assessing a publication? Can you tell that full data is only being pulled from Thomson Scientific's Journal Citation Reports?

Impact factor calculated as total number of times articles in the journal, from the past two years, were cited in the index of the current year of Journal Citation Reports (ie. Number of citations), divided by total number of citable items in that journal (i.e, for 2012 index we'd be looking at 2010 and 2011). Comparative measure of journals in the field. Thought to be gaming in how the denominator is calculated, what is considered citable.
A post by Todd Carpenter on the Scholarly Kitchen, the blog maintained through the Society for Scholarly Publishing, praised the development of indicators like the H-index and the Eigenfactor.
Carpenter writes:
"Fortunately, tools are quickly falling into place to provide metrics on the things that the larger scholarly community really needs: individual items and individual scholars. The broader scholarly community has been aware of these needs for some time, and we are making considerable progress on implementation. There has been

a great deal of progress in the past decade in increasing the granularity of assessment metrics, such as the h-index and the Eigenfactor. These metrics are gaining traction because they focus on the impact of a particular researcher or on particular articles, but they are still limited by dependence solely on citations."

Slide 6.
Impact factor has come under fire recently—here are two pieces of criticism, one from the altmetrics manifesto, another published in the Chronicle of Higher Ed. The Chronicle piece is focusing on DORA, the San Francisco Declaration on Research Assessment.

Slide 7.
Meeting of editors and publishers of scholarly journals in December 2012, during the Annual Meeting of The American Society for Cell Biology in San Francisco. Recommendations, known as the *San Francisco Declaration on Research Assessment*, was developed, and they are collecting signatures now. This effort has issued recommendations for funding agencies, etc., in evaluating impact using metrics.

Quotes from DORA [bold added]:
"The outputs from scientific research are many and varied, including: research articles reporting new knowledge, data, reagents, and software; intellectual property; and highly trained young scientists. Funding agencies, institutions that employ scientists, and scientists themselves, all have a desire, and a need, to assess the quality and impact of scientific outputs. It is thus imperative that scientific output is measured accurately and evaluated wisely. The Journal Impact Factor is frequently used as the primary parameter with which to compare the scientific output of individuals and institutions. **The Journal Impact Factor, as calculated by Thomson Reuters, was originally created as a tool to help librarians identify journals to purchase, not as a measure of the scientific quality of research in an article. With that in mind, it is critical to understand that the Journal Impact Factor has a number of well-documented deficiencies as a tool for research assessment.** These limitations include: A) citation distributions within journals are highly skewed [1–3]; B) the properties of the Journal Impact Factor are field-specific: it is a composite of multiple, highly diverse article types, including primary research papers and reviews [1, 4]; C) Journal Impact Factors can be manipulated (or "gamed") by editorial policy [5]; and D) data used to calculate the Journal Impact Factors are neither transparent nor openly available to the public [4, 6, 7].
Below we make a number of recommendations for improving the way in which the quality of research output is evaluated. **Outputs other than research articles will grow in importance in assessing research effectiveness in the future, but the peer-reviewed research paper will remain a central research output that informs research assessment. Our recommendations therefore focus primarily on practices relating to research articles published in peer-reviewed journals but can and should be extended by recognizing additional products, such as datasets, as important research outputs.** These

recommendations are aimed at funding agencies, academic institutions, journals, organizations that supply metrics, and individual researchers."

Slide 8.
Value of single number is limited and flawed. But some of the issue with the impact factor comes in how it's been misapplied.

Slide 9.
The altmetrics manifesto, from 2011, set forth a vision of how altmetrics could enable a new filtering of scholarly resources. It took as its aim both resources that had not been traditionally produced or accounted for– (1) datasets, slide decks, student works, websites, digital projects– (2) and works such as journal articles that, while being standard academic output, were suffering from a lack of post-publication filtration. The altmetrics manifesto frames existing practices of storing, citing, using, reviewing, mentioning resources, by scholars and those outside of academia, as central to a new mode that can electronically account for these practices and thus aggregate and make that information available to the researcher, the institution, the general public.

Slide 10.
The manifesto is, of course, also framed as a critique of the traditional, "swamped" system of assessing and filtering scholarly materials, which it saw as overly reliant on slow, outmoded systems that provided failed proxies for impact. In particular, it criticized these three: peer review, citation counts, and the biggest enemy of all, the journal impact factor.

Slide 11.
Jason Priem, who is something of an altmetrics evangelist, was featured in a NISO webinar recently, speaking about the development of this crowdsourced peer review, so rather than being a closed, blind peer review, a process both central to many traditional modes of scholarly communication and criticized for being "review by two," and for, in certain fields especially, being a biased, conservative, slow, resistant, and self-replicating process.

Galligan and Dyas-Correia, in a 2013 *Serials Review* piece, provide several definitions for Altmetrics and point to their commonality: "altmetrics examine the content of the social Web in order to provide either an alternative or enhancement to the use of journal impact factors and click-through rate analysis to measure the impact and value of scholarly work."

Slide 12.
So here I want to be sure to make a distinction between various metrics. Some, like Google Analytics or Springshare Analytics, which we'll look at in a moment, are not explicitly altmetrics but fall into that larger bucket of webometrics, assessing usage alone, page views, download count.

The impact factor is a journal-level measurement. Citation count is likely at the article or book level. The Hirsch index is another bibliometric that has attempted to break down some of the monolithic nature of the impact factor. It looks at citation, but at the author/researcher level, at the researchers' publications, rather than the container of publication, i.e., the journal.

Article-level metrics are described in a recent SPARC primer as "similar but not synonymous" with altmetrics. Article-level metrics "are an attempt to measure impact at the article level. In doing so, ALMs draw from a variety of different data sources, some traditional (e.g. times cited) and some new (e.g. tweets). The attempt to incorporate new data sources to measure the impact of something, whether that something is an article or a journal or an individual scholar, is what defines altmetrics. Altmetrics are about the data sources, not the level of aggregation. ALMs are about the incorporation of altmetrics and traditional data points to define impact *at the article level*."

If you're a little confused about the distinction between these metrics and "alt" or "alternative" metrics, you are not alone. NISO just released a draft whitepaper on alternative metrics, which I highly recommend, which recognizes that we're in something of a wild west of alternative metrics claims and applications. It lays out recommendations for places where NISO might intervene. For example, in defining the scope of what "altmetrics" means, or suggesting a different terms, or establishing subcategories for altmetrics.

Slide 13.
We use metrics in the library all the time, they help us document usage, page views, etc. Our OAKTrust repository's built-in statistics package shows the number of item-level and bitstream-level visits to a file over time, as well as the country/city views. Springshare provides statistics on our LibGuide usage (thanks to Melissa Edwards for sharing that with me). We had integrated Altmetrics into our Primo search, though it's now been taken out.

Increasingly, I think, there's the expectation that digital sources will provide use with usage information. And for websites, LibGuides, items deposited in the IR, these items might never have been covered by traditional metrics because they didn't exist before. Metrics allow us to point to and count these resources in new ways.

In the sciences, you can see this ability to count and assess research output that isn't appearing in journals—datasets, code, blogs, posters, lectures and online class or teaching materials, ETDs, etc.

Slide 14.
Examples of where alternative metrics are being collected and exposed.

Slide 15.

A growing body of literature is attempting to discern overlaps between altmetrics and citation analysis– there's this question of whether altmetrics are predictors of citations, such that certain measures picked up by altmetrics might indicate a likelihood that an item will receive a certain number of citations. Which has some logic to it-- If I'm saving articles and datasets in a Mendeley folder, is that a predictor of my citing it (or of someone else citing it)? This literature is thus far inconclusive, to the best of my reading of it. But there's also been this thread that looks at "flavors" and attempts, basically, to classify varieties of usage and impact, and to characterize scholarly items along those lines. For example, articles that are widely read by scholars but never make it over to the public. Or are found on scholarly sharing networks but not yet well-cited.

Slide 16.
This is a demo of Altmetric—I want to clarify here that Altmetric is a company and tool that works in the altmetrics arena—very confusing naming-- hooked up to PLoS, showing the articles mentioned in the past minute on social media. So the service is letting you discover and see "hot content."

This is an old screenshot. But you'll notice that there might be something of a skew in terms of what's hot. You can see the top-tweeted paper is "An Introduction to Social Media for Scientists."

I'd argue that this gives us a fuller sense of scholarly communication as it unfolds, at least in terms of open scholarly communication. Micah Vandegrift at FSU points to Thomas Kuhn's statement that "Spotting emerging research fronts will require tracking 'formal and information communication.'" Altmetrics and webometrics align with this prediction. Not just a static measure, determined at the moment an article is accepted for publication.

Stacy Konkiel, formerly a librarian at Indiana, now based at ImpactStory, accedes that "Altmetrics are not currently as user-friendly as the journal impact factor. Critics note that in lacking a single number, rating, or score, altmetrics require scrutiny and interpretation that can be burdensome to end users." But they also offer context, with services linking out to the tweeted mentions, Mendeley saves, etc. And the fact that altmetrics aren't appearing as single, again monolithic scores, has value in forcing this more contextual evaluation.

Slide 17.
Altmetrics, of course, are some of the only options available for systematically tacking the influence of products like slidesets or websites.

Galligan and Dyas-Correia elaborate further on their vision of how altmetrics and bibliometrics will interact. They write: "Initially, altmetrics will compliment traditional metrics, removing the complete reliance on the impact factor. They will provide wider and deeper insights into the impacts of scholarly articles, researchers, departments, universities, and so on– outside the reach of traditional metrics, which

rarely exit the bounds of 'traditional' scholarship. This is where outreach again becomes a key driver. Many fields, but particularly the sciences, are increasingly pressured to demonstrate how research is relevant to the general public. It is important to combat an environment of suspicion and propaganda, painting researchers as working on ineffective products funded by taxpayer money. Demonstrating relevance and utility is therefore essential to sustain a culture where research is encouraged and supported. When this relevance is measured and quantified, the system functions more efficiently and provides more evidence to support research work. In the long term, altmetrics may perhaps outlive the impact factor and other traditional citation-based metrics, but that would most likely be dependent on a substantial shift in the way scholarly output is presented in its most formal state."

Slide 18.
So, one challenge is extracting the data and ensuring that the data is openly available, has an open API, and so can be collected and analyzed. Metrics may not stay permanently available even if they are available at the moment. Cited throughout the literature on altmetrics is the scholarly social bookmarking tool Mendeley, which is much larger even than CiteULike, its next-largest competitor. But Mendeley, of course, has been purchased by Elsevier. As David Dobbs wrote in the *New Yorker* at the time: "Elsevier has two reasons to buy Mendeley. One is to squash it—to destroy or coöpt an open-science icon that threatens its business model. Many critics fear that's the case. The other reason is to possess the aggregated data that Mendeley's users generate with all of their searching and sharing. Mendeley is still growing, with two million three hundred thousand users sifting through over a hundred million references. Their use patterns reveal who is reading what, which papers are popular, what lines of research are surging, which disciplines and journals are crucial, and a lot of other extremely valuable information.
"... Now Elsevier owns all of that data. But if it wants users to continue generating streams of data, the company will have to play nice, which leaves it with something like the Facebook model: create software and a huge social network in which people share information that it can profitably harvest, and be just conciliatory enough about privacy, anyway, to repel fewer people than it attracts."

And, if you remember, when Elsevier purchased Mendeley, many users very publicly declared that they weren't going to continue using the product, and contributing to an income stream for Elsevier.

There's also a lack of metrics available for some items: it's hard to get much in the way of usage information about book chapters. Metrics are dependent on a web infrastructure that demands persistent, unique, disambiguated products. ORCID running on CrossRef. ImpactStory requiring DOI, GitHub, SlideShare, etc. There may also be versioning issues present.

Slide 19.

Galligan and Dyas-Correira write: "Altmetrics have always been closely associated with open access publishing." They point to the shared goals of openness between altmetrics and the OA movement and refer to Heather Piwowar, of Impact Story's, post on altmetrics in service of "open impact tracking." Altmetrics is somewhat geared to prove the value proposition of open access, with open research and data having a greater potential for public and even scholarly impact, as demonstrated through altmetrics.

Both altmetrics and CC licenses are built into open science, really, as I'm hoping this workshop demonstrates. They are two elements of a larger infrastructure that is being constructed and adapted right now, to support scholarship that adheres to the principles of Open Science. It's an infrastructure necessary as we move away from a concept of a field and its output as organized into journals with distinct tiers of prestige, that are behind paywalls and that own published scholarship.

As we're seeing data, code, teaching materials, blogs, and other outputs distributed, we're seeing more need for these systems.

Creative Commons is really a cornerstone of the OA movement, particularly as OA is defined as "libre" and not just "gratis"—that is, not just items that are available without payment. But items that are available without certain rights restrictions, whose reuse is specified.

I'm going to play a short clip that I think gives a strong overview of CC licenses now: http://creativecommons.org/videos/creative-commons-kiwi

Slide 20.
The open access movement, as you're likely all well aware, is complex, multifaceted, both international and very local in application, and longstanding. Creative commons licenses are just one aspect of the movement, but I want to spend a little time on them now.
How many of you are familiar with CC licenses?
How many of you have included CC-licensed content in your own work or performed searches of CC-licensed materials, say, looking for an image to include in a powerpoint?
How many of you have applied CC licenses to your own work?
And who has set up licensing for larger groups or projects and relied on CC licenses there?
We're grown somewhat accustomed to CC licenses now, and I think it's worth reflecting on how remarkable they really are. The group behind them was founded in 2001, with support from Duke's Center for the Public Domain, and released the first set of copyright licenses to the public in 2002. According to the institutional history on the CC website, these licenses were inspired in part by the Free Software Foundation's GNU General Public License. By 2003, about 1 million licenses were in use. The current version of the licenses, 4.0, is designed for international use. There is a core suite of 6 license; additionally, there's CC0, which is outside of the suite in

that it doesn't require attribution– it's essentially a license that grants content to the public domain.

Slide 21.
There are options for licensing work beyond CC. You may have seen the DPLA's Getting it Right on Rights initiative, which just received Knight Foundation funding, and is aimed at building out digital library licenses to accommodate things like unclear rights status and other international rights issues– the funding will also cover Europeana and other non-US massive digital library projects. They are aiming to develop a new licensing namespace that can plug into the CC infrastructure and license set. There are also initiatives like Open Data Commons that are aimed at providing licenses more specifically tailored to data and databases.

Slide 22.
Instructions for exercises (distributed)

Slide 23.
NISO Altmetrics Standards Project Whitepaper—encourage feedback