With Great Power Comes the Responsible Use of Metrics

Title slide

Welcome to this webinar on the responsible use of metrics. As comic book fans will know, the title is a slight play on a quite from Spiderman but sadly for any fans this is as far as I'm taking that theme!

Topics

Whilst I'd like to spend our time talking about superheroes we're here to talk about metrics and how to measure research in a responsible way. To do that we'll be covering the following topics:

- The current system (and why it needs to change)
- One possible solution in the form of the responsible metrics movement
- A closer look at why responsible metrics are important and what is actually being done to support their use
- And how library staff can work to support their research community with responsible metrics

Metrics today

Before we look at responsible metrics in any depth it's important to know where all this is coming from so we're going to start with a quick recap on metrics and the current situation.

Types of metric

No doubt you will know that metrics are numerical measures which are used to assess the quality and impact of research. In the context of research support there are two main types of metric that library staff are likely to be concerned with: biblio and Almetrics.

Bibliometrics are the ones most people will have heard of. These involve the statistical analysis of publications and other outputs and include metrics like citation counts, the H-Index and the Journal Impact Factor. A lot of these metrics are entrenched in academia and it can be very hard to get researchers and academics to let go! They are quantitative measures and so very focused on the numbers without really looking at anything else so many people have argued that they can be misleading.

Altmetrics were created in 2010 as a way to counter some of the problems of traditional metrics. Perhaps the biggest problem is that traditional metrics were designed for a world dominated by print publications and it was becoming increasingly difficult to measure newer types of output like social media posts. Altmetrics look at a range of sources used to share research and gives them a different weighting depending on its potential impact – so a news story would have a higher weight than a social media post. The result is a number and something known as the Altmetric doughnut which is a colour coded wheel showing where an output has been mentioned. Altmetrics also allow you to look at what has actually been said which is something traditional bibmetrics don't do.

Levels of metric

Between these two categories there are lots of different metrics which are used to measure different levels. The model on the screen is from a book called *Meaningful Metrics* by Roemer and Borchardt. It's a useful model as it highlights the different levels of measurement:

- The first level is the *Individual scholarly contribution* – the output of a research project such as a journal article or a book chapter. Obviously the range of output options has increased dramatically in recent years and authors can produce a range of contributions from a single project, each with it's own metrics attached.
- The *Venues of production* – refers to the title or the format of the publication, for example the journal title that an article is published in. Researchers will often consider things like the impact factor of a particular journal when they are trying to decide where to publish their work and it is often a major part of their decision making (rightly or wrongly).
- Of course *Individual authors* – are also measured and how they perform is often taken into account when institutions are looking at things like promotion and tenure. There are a whole range of metrics dedicated to measuring authors at different stages of their career and metrics for individual authors can vary widely depending on the calculations used.
- The final level is *Groups and institutions* - which obviously measures the impact of the wider institution a researcher belongs to. This might be measured in terms of how many researchers publish in high impact journals, how often people from a certain department are cited by peers or the ranking of one university compared to another.

This is a slightly silly example but it helps to illustrate the point. Dr. Snickers of Candy University has written a paper published in the Journal of Chocolate Studies. Each of these four levels of metrics can be assessed on its own or combined as needed to present a picture of impact whether you are trying to measure the paper, the journal, Dr. Snickers or the university as a whole.

Uses of metrics

It should be obvious form this that metrics are an important part of the research lifecycle. The implications of measurement are present at every point – researchers will have to outline the potential impact of a project when filling in their grant application at the start and will need to add metrics on the outputs to the CV at the end. On the screen you can see a summary of some of the common uses of metrics in academia:

- Judging the quality of research – Rightly or wrongly, metrics are often used as a substitute to judge the quality of research so if it has more citations it's better research (no matter what the content of these citations is). We'll look at why this is a problem during the webinar.
- Publishing decisions – Researchers often rely on the metrics of a publisher or journal title to decide where they will publish their findings. They will want to publish in the titles which have the biggest impact as they think this will be beneficial to their careers and help them to get more work and respect within their discipline.

- Performance reviews – Many institutions use metrics as the basis of performance reviews and benchmarking exercises for individual researchers, departments and the whole institution. They will see how often the outputs of a department are cited or how many of their staff are published in certain journals. There have been stories of some institutions providing *encouragement* to their researchers to publish in certain titles – for example the journal Nature is highly sought after as one of the highest impact journals in the sciences and there have been claims that some institutions will pay up to $20,000 to the lead author of any paper published there.
- Career advancement – Linked to this is the use of metrics to advance a researcher's career. They are often consulted as part of the hiring process and the better the metrics someone has achieved, the more likely they will be considered for a job or a promotion. In the US in particular metrics are heavily emphasised during the tenure process and the more impact someone has had, the more likely they will get a secure academic position.

Limitations of metrics

There are obviously major decisions which are being made on the basis of metrics but should this be the case? Metrics as a measurement tool obviously have their uses but they also have many limitations:

- Quantitative measures – Most of the metrics in common use are quantitative measures which means they put a lot of emphasis on numbers over any other measure of impact. Of course, numbers tell part of the story but it is important to look at other impacts. A good example of this is the Andrew Wakefield paper from the 1990s which linked autism with the MMR jab. Although the research has since been widely discredited the paper still has a really high citation rate as people often cite it as an example of poor research. Looking at the numbers alone people might assume that this was a good paper and this shows why it's important to look beyond the numbers to the quality of the work or the researcher. Altmetrics have gone some way towards addressing this problem but there is still work to be done.
- Potential bias – No one metric can capture every mention of what it is trying to measure so they all contain some form of bias. Some of the most well-known metrics are owned by commercial companies who will base the calculations only on databases they own which obviously don't contain every mention of a work. It can be really hard for researchers to verify the calculations behind a score and make sure that it's accurate – especially if it's based on proprietary software or a database that they don't have access to.
- Lack of consistency – Linked to this is the lack of consistency between metrics. A researcher can get different scores using different measures as they are all based on slightly different sources. On the flip side of this, a lot of metrics fail to take into account the differences between disciplines and career stage. Metrics which count things like citations and publication numbers obviously favour those at a later career stage as they have had more time to build these up and some early career researchers feel disadvantaged by the most popular metrics. There is also something known as the Matthew effect where established researchers are cited because of their reputation – people expect to see certain works or people mentioned in literature reviews. This increases their metrics which only enhances their reputation further and on it goes.
- Potential for gaming – As well as genuine differences, it's important to acknowledge that there is the potential for gaming the system. Metrics can be quite easily manipulated and although the process is frowned upon that doesn't mean it doesn't happen. Authors can self-cite their work, or get their colleagues to do it, people within an institution can arrange to cite each other and there are also problems like gift authorship where the author of an output will add their supervisors name to it to enhance both the output and the reputation of their supervisor.

- Fit for purpose? - Perhaps one of the biggest problems with metrics is that they were not intended to be used in the way they currently are. Traditional bibliometrics were originally devised as a way for librarians to select which titles they wanted to stock in their library – obviously if something was cited a lot by researchers in a department it would be a good idea to have it in stock. Given the major decisions that are now being based on these metrics many people have said that they are not fit for purpose. Added to this is the fact that many metrics have been in use for some time and haven't really adapted to changing methods of publication, despite the introduction of Altmetrics.

Responsible Metrics movement

One potential solution to this problem is the move towards the responsible use of metrics in assessment. Rather than calling for them to be replaced or cut out altogether, the movement advocates that they are used in a more responsible way including placing less of an emphasis on numbers and taking into account a wider picture of impact and the facts behind the numbers.

Responsible Metrics is not one single set of principles or rules but a general move towards making sure that assessment is fair and robust. Several different groups and documents have contributed to the development of responsible metrics and I'm going to spend some time looking at three of the most important - DORA, the Leiden Manifesto and the Metric Tide report – before I sum up the main points.

DORA

DORA – the San Francisco Declaration on Research Assessment – was developed in 2012 following conversations at an American Society for Cell Biology meeting after attendees got talking and found that they were concerned with the overreliance on metrics to assess their work. They were especially concerned that their work was being judged more on where it was published than on its quality – for example the same paper would be judged as more important if it was published in a high-ranking journal, even if the content was the same. Together they came up with the principles of DORA with different levels aimed at organisations, individuals and publishers. Although this was developed by those in the sciences it's important to stress that the original group aimed for the principles to cover all disciplines.

DORA is a document which anyone – individuals or their wider institutions – can sign up to to demonstrate their support for the principles of responsible metrics. To date over 14,000 individuals and 13,00 organisations have signed including all seven of the UK research councils.

Leiden Manifesto

Next to be launched was the Leiden Manifesto for Research Metrics which was, like DORA, conceived at a conference by a group of social scientists and research administrators. The Manifesto was released in 2015 in an article in Nature which outlined both the rationale and the ten key principles. Like DORA, the Manifesto has some high-profile adopters from around the globe. The principles go into detail on how both researchers and their outputs should be measured as well as stressing the importance of verifiable data – a key component of the move towards open research. The Manifesto is currently under review to take into account contributions and discussions from the wider research community.

The Metric Tide

Released in 2015, The Metric Tide was an independent report commissioned by HEFCE (the now defunct Higher Education Council for England). The report was produced by a group of researchers from various disciplines and was designed to investigate the reliance on metrics ahead of the next REF (Research Excellence Framework). The report advocated five key principles for responsible metrics and concluded that the general concern about metrics was justified and that any future REFs should take into account qualitative measures of assessment to balance out any numerical measures of impact.

Responsible metrics

There isn't time to go through everything in detail in this webinar but there are some themes which are common to all of the documents and the Responsible Metrics Movement as a whole:

- Qualitative and quantitative – Responsible metrics advocates a mixture of both types of impact to produce a more rounded picture of influence. Although numeric measures have a place, they should be complimented with wider measures such as mentions of the research in the popular press. Tools such as Altmetrics have gone some way towards this but there is still more work to be done.
- Openness – Linked to the general move towards open research, the methods used to calculate different metrics should be open and available. This enables researchers to audit metrics and better understand how they are devised. Crucially this means that any measures used are verifiable which is a vital part of ensuring research integrity.
- Quality – A key concern from researchers is that the metrics achieved by a particular piece of research or title are becoming more important than the findings themselves. With authors under pressure to publish in titles with the highest metrics they are concerned that the merits of the actual research are being overlooked. There is also a need to consider the global nature of research and ensure that research being done outside of the global North is judged equally on its merits.
- Range – It is important to consider a range of measures of any one researcher, institution or output as relying on only one metric can give a distorted picture of impact. Different measures will present different results for the same area which can be confusing. Certain metrics are often popular in particular disciplines and again this can create problems. Finally it is important to consider a range of metrics to avoid relying on those which are biased towards researchers at a more advanced career stage.
- Review – Any metrics used should be regularly reviewed and updated to ensure that they are still fit for purpose. This is also a chance to adapt them to take into account developments in scholarly publication, something not all existing metrics have managed to achieve.

With the increasing emphasis on metrics and measurement, issues like these are going to become more important. Different researchers and disciplines will be impacted in different ways but since assessment underpins almost all parts of the research lifecycle it's safe to say that all researchers will come across it at some point.

Supporting researchers with responsible metrics

So I thought it would be a good idea to end the webinar by talking about what librarians can do to support their research community in this area and how you can advocate for responsible metrics. A lot of this will depend on factors like discipline, career stage and how things operate locally for you but there are some general principles:

- Practical support – On a practical level you can advise researchers to try to make sure that all of the work is linked back to a single, authoritative source like the published version or the one in a repository as this will make keeping track of metrics much easier.
- Advice – Working in a library it is likely that you'll be asked about some aspect of metrics at some point. You don't need to know how each one is calculated but a vague awareness of what they are and how they work is a good idea as well as some knowledge of areas like Altmetrics. Once you have a chance to talk to people about metrics you can move on to advocacy.
- Advocacy – I would recommend trying some stealth advocacy. Rather than running a whole session on responsible metrics like this one I would use any opportunity to highlight responsible metrics, perhaps in the context of a wider conversation about impact and sharing work. It's important to advocate up as well as down – talking to researchers is great as they can increase the pressure to change the system but educating those with decision making power can have a greater influence – let them know that decisions should be based on more than just metrics.
- Sign your name – Organisations and individuals can sign any of the declarations we've discussed (*and several Cambridge departments and individuals have already signed DORA*). If you want to show your support for the principles of responsible metrics you could consider signing and making that public commitment. The more people who sign, the stronger the movement will grow.

It's important to remember that responsible metrics is still at an early stage in the UK and other countries like the US are far more advanced. Perhaps the most important thing you can do to help support your researchers is to keep abreast of developments and learn more about what the responsible metrics movement is and how libraries can support it.

Research Support Handy Guide

To help you do this we've produced another of our Research Support Handy Guides on responsible metrics. These are guides aimed at helping librarians understand some of the key issues in research support but you can also print them out to give to researchers or adapt them if that's easier.

Thanks for listening and watch out for more Wednesday Webinars coming soon!