Transcript for Australian Scholarly Summit Panel:
Research Metrics and Recognition

Anthony Dona (Clarivate Analytics):

Thanks very much Thijs. I just want to thank Taylor & Francis for the opportunity, I presented at Monday's summit in Sydney and they've invited me back, which is always a good sign.

I'll introduce myself first, I'm an academic by background, I did a PhD at the University of Sydney and moved over to Imperial College and did a post doc there, before coming back to the University of Sydney to continue my post doc, and a couple of years ago I moved over to what was then the Science IP sector of Thomson Reuters, which has since broken away. So in the middle of 2016 the Science IP sector of Thomson Reuters broke away and was sold off, and has rebranded in the time since and is now Clarivate Analytics.

So largely now my main focus and my day-to-day interests are on understanding the impact of science and how we actually hone that and nurture what is very, very good science. I do still hold an honorary position through the University of Sydney and I still do review and I still do publish actively with my collaborations there, I just unfortunately don't get paid for that activity any more. So like I say, my day-to-day interests have changed a little bit.

This you'll see on a lot of slides with people that kind of do the sort of things that we do. Underneath us you'll find what our platform's for, the discovery of science and the understanding of the discovery of science, through to the protection and the commercialisation of science, so I do want to kind of clarify that. If you look at our business it's broken up into three main sections, and I myself work for what is the 'government academics' section. So I work with universities across Australia, New Zealand and the Pacific, which is my region. I am not a sales person, I don't sell to these universities, I simply help them understand our solutions and how they can use them.

You will know and probably remember ISI, which became the Web of Science, which recently was announced is going back to ISI. We recently had Annette Thomas move over to our company who used to head up Springer Nature, and she has a vision where we have an academic group at the center of our government academic business, a bibliometric group, a group that really looks to measure the impact of science, and then we build our platforms off of that.

I don't know whether any of you know our other platforms in this sector? There's Journal Citation Reports, which is where our Impact Factors come from, and we have Insights, which is our bibliometric analysis tool, which universities use, almost all Australian universities use them to measure their own impact, and it really is an analytics tool to understand where they're

collaborating, where they're having impact, which journals they're operating in and a whole bunch of different kind of explorations and reports they can produce.

The other two businesses that I won't talk about too much are our life sciences business, which is largely clinical trial databases and what is drug discovery. We've seen that now moving into the government space, we're seeing universities very interested in this as well as what was corporations that we were working with, as well as intellectual properties. So the third business there is intellectual property, which is what was Thomson Innovation and has now been rebranded to Derwent Innovation, and it's our patent and IP database. Which again is now also moving into the academic space as well, we're getting innovation officers in universities and a lot of universities doing that very, very well. So we're seeing that shifting into the academic space as well, which is really exciting for someone in my role because I get to kind of cover all three businesses.

Now these are the questions that we'll get from editors, which is why I wanted to address this on this slide. We have very strict criteria when it comes to the Web of Science as to what journals we index and which journals we don't, and which journals are getting measured in terms of getting Impact Factors and which don't.

So these are the criteria in bullet point form. We have a set of editors that are independent from any publisher that actually pick and choose which journal should be going. And the main things we're looking for there are really a timeliness of publication, which is an interesting topic of conversation. I think one of the slides over lunch kind of addresses that. We also want to see a transparent peer review process, an editorial process obviously. And we also like to see international recognition of a journal, so a journal that has interest in citation networks that don't just connect within themselves, but connect to what is an international audience.

Now this was one of the criteria that we were constantly failed on by a lot of people, because a lot of journals and a lot of subject areas actually don't get international recognition for very, very good reasons. If you're operating in business or economics or law or even certain medical areas, you'll actually find that they don't apply internationally. Even indigenous studies within Australia probably doesn't apply so much to the international audience. So what we've recently done and recently released is the Emerging Sources Citation Index. And we've also already chosen seven thousand journals that we're indexing in that, which is added to the three already kind of established indices within our Web of Science platform.

We also developed that and fast-tracked a lot of that for the Australian market, because we're the citation provider for what is the Excellence in Research Australia (ERA) process. So the ARC's ERA process, we went to tender for that to provide the data for that, and we won that, so we're providing the citation data across what is the ARC's journal list for that. And I was heavily involved in that process, so if there are any questions around that, I'm quite happy to field them.

So that's that. And I guess the dot points in green there are the criteria that we have for the

Emerging Sources Citation Index, and the rest of the criteria are things that we're looking at for our three hard indexes, which are our Social Sciences, Science Citation Index, and Arts and Humanities Index.

I wanted to talk about something kind of exciting as well within our business. This morning, I happened to be in Melbourne anyway, there's a Publons breakfast going on at the VALA conference, so I'm here for the VALA conference working with librarians across universities and the kind of government space. So we've attended that and we had this morning a really very nice breakfast, not quite as nice as this room, but really very nice where we presented Publons.

Now has anyone heard of Publons? Yeah, cool. So it's really, and probably as editors yourself you know this all too well, and I know this from a reviewer's perspective, in that the only reason I would ever review an article when it was sent to me was out of the goodness of my little heart for the area of science that I work in, and also kind of to get a heads up on the paper that was being published. I'd look at the authors, I'd probably know the authors because I know the space that I operate in and I'd want to read their paper before anyone else got to read it.

So that was why I was doing it, and that was why I was engaging in that process. Now there's no recognition and it's very difficult to get recognition for peer review because it's all blinded, as well as a lot of the time the editorial process isn't recognised as heavily as it should be.

So this is a really cool story, two guys from new Zealand who are best mates growing up, both went to Victoria University of Wellington, and we had the University librarian at the UW speak this morning on this process. Basically what they do is they've started to work with publishers and capture what is peer review. So they're capturing who and when people are peer reviewing, and who and when people are editing, which is really very cool. Now obviously the process is still behind closed doors, but you'll get a tick next to your name every time you review for a journal. They're capturing this information in two ways, one is working with publishers themselves, so Taylor & Francis is of the 12 biggest publishers they've already signed up. And I think there's 400 odd journals that are already capturing … from Taylor & Francis that are capturing information that way and they plan on expanding that out to many, many more of their journals in the future. And maybe you can correct me on that exactly how that process is going to go, but there's also, yeah, nine other of the big publishers, and they've signed on many, many, many publishers that you would have heard of to that process already.

The other way is that reviewers can actually send the thank you email to Publons itself, and get recognition for it that way, so they can go through this manual process. This is what a typical kind of profile looks like, and we've already had about 300,000 reviewers already signed up. So 300,000 academics have already signed up and created a profile. And about 10,000 each week are signing up at the moment, so it's pretty fast growing, somewhere between five and 10,000 reviewers each week are signing up to Publons.

So here you can see Joanna, she has a wonderful little profile - she edits for a number of journals and she's reviewed 18 times for PLOS Computational Biology, seven times for PLOS

Genetics, and so on and so forth. So this is all public, and if you go to Publons.com you can filter through this information and see who's already on there.

Now where this grows, and where this expands, when the boys started this, they actually wanted to see transparency in the peer review process (and with a lot of brick walls that actually doesn't work so well because, well, there is genuinely a lot of brick walls) to show a completely transparent peer review process, and that's a very interesting conversation. However, this data is actually very, very useful in many ways for not only publishers, but also institutions, also government and funding agencies. So they're now capturing this, and producing dashboards. So one very interesting thing, and as editors I'm going to talk to this one very specific point that has been aided through this process, thank you.

So 52% of invitations by editors get accepted by reviewers, potential reviewers, so it's about 50% right? So one thing that editors spend a lot of time doing and wasting their time on is finding a reviewer in that specific space that is willing to review. Now what Publons does is obviously captures people who are reviewing. We're with Publons, we acquired Publons middle of last year, and we're now working together to combine our datasets, so not only our published literature datasets, but our reviewing where people are reviewing datasets. And they've generated this wonderful tool, which actually looks at a person or an academic in a 360 kind of view of them. So publishing, reviewing, editing, and where they sit in that kind of space.

I'm going to very quickly show you this, this is a tool that is in development, and we're very much welcoming comments. It's called Reviewer Connect, you get a publication sent to you that's been sent through for review. You put the title, the abstract, the authors in if you want, it looks for conflicts of interest between authors, and authors that may have already authored with the authors that are on this paper. You can also exclude reviewers manually from what's shown here. And there's kind of what it will look like filled out, and what this does is churns through Publons data as well as Web of Science data to pull out a list of reviewers. If we have email addresses, we obviously have contact details for a lot of people within our database, the email will pop up. If they have a Publons profile, that will pop up as well, or an author ID or what have you. And also keywords that they've either put in or have been pulled from the title abstract that you've kind of entered there.

You can open up one of these and have a look at the very basics on one of these people's Publons profile where they're affiliated, you know, the keywords, and where they've reviewed in the past. So this is a kind of, you know, have they reviewed already 10 papers this year, or have they not done anything this year?

So I guess it's one of the tools that we've seen that could be targeted to editors as a timesaving tool, which is why I wanted to highlight that one today.

Good, I'll stop there because I have dinged. I do welcome questions, I will be around for lunch and I do welcome kind of conversational questions and people coming up and disturbing me. Thank you.

4

<u>Leonie Hayes (Digital Science – Altmetric):</u>

Okay, thank you very much. I'm Leonie Hayes, I'm from Digital Science Australia, and most of my working life I actually sit behind a computer. I've got a background as a librarian, and a lot of work on system integrations and system administrations, so it's lovely to be here talking to the real creators of content. I always like talking to academics.

So I work, as I said, for Digital Science, and Digital Science is a larger company that has portfolio companies, and each of those portfolio companies has a sort of focus on a particular area within the research life cycle. So these companies are startup companies, most of them starting in academia, and so those founders of those companies get invested in by Digital Science and most of them stay with the company through their life. So the product that I'm going to talk about is Altmetric. I wear many hats in Australia, because we're quite small. I look after Altmetric, Symplectic Elements, and the new Dimensions product we've just launched last week.

We aim to support researchers at each of those life cycle points, and Altmetric is looking for attention and measuring that in different ways.

So the Altmetric story is not very old, it's a fairly new company, started in 2011 by Euan Adie, he was working at a university in the UK, developed this cool database, and then in 2012 they launched an Explorer for the publisher's platform. We get all the data from the publishers, so our relationship with the publisher underpins the database, so that's really important. Good quality metadata from publishers is the key to making that database work.

During 2014, they had the first Altmetric conference, and then in 2015 they launched a product, so there was a lot of comments around, you know, that's great that you're supporting mentions for journals and conference papers, but what about books? For a lot of academics in arts and humanities, the book is their main item. So Bookmetrix was a database that looks for attention to books.

In 2016 the Badges for Books, and they relaunched the Explorer product for institutions, so many of you will actually have access to the institutional product which will give you a different type of access, and they included a lot of additional functionality. And then in 2017, we've got around 230 customers, there's 18 Australian universities who use the institutional product, a couple in Asia and some in New Zealand as well.

So I thought for a bit of fun, that I would use the Altmetric database today and look for things that had the title 'Valentine' because it's Valentine's day-to-day. And that's a frivolous topic a lot of the time, but actually when you look at this particular article here, in the conversation it's a very serious article looking at roses and it gives you some clues around sustainable shopping. So I just thought that might be nice. If any of you haven't seen the Explorer product, that's what it looks like. You get a list of those colourful donuts with a score, which is not the main

focus. So it's the colours within the donut and the type of attention and what's been said.

So why Altmetrics? And the reason is because they complement traditional citations, so people who are new and emerging researchers who've recently published their PhD thesis, who haven't published a lot of journals, or yet to publish, might have some datasets or conference papers. And that picks up attention to them. So we look for mentions in news reports, references in policy documents, and the policy document area is something that's growing all the time. And later on this year we're going to be adding patents as well.

So this is a timeline of attention which might happen. And somebody mentioned the other day the difference between normal Twitter activity and academic Twitter. So academic Twitter now is a huge movement, and Anthony mentioned on Monday that there's some, well, there's a few different arguments around this. But there is a little bit of evidence to show that there is a correlation between early Tweets (especially academic Twitter) and future citations.

So over a time span, what you can see along the bottom there is in hours, days and months. And you can see towards the end of that scale is going to be where you're getting citations. So when you've published your research online, if you can grab the attention to that early on, we track that within our database.

We look for things in unexpected places. A lot of the engagement and impact in Australia is looking for things of a more scholarly nature, so things like post-publication peer review in the F1000, policy documents, news, particularly sort of scholarly articles like The Conversation articles, Wikipedia, those types of things go to make up that sort of space now. So people would say, oh, that's just, you know, Altmetric, that's just about Twitter and Facebook, so it's way more now.

So what do we need to track the attention? We need an output, it can be a journal article, the dataset, a non-traditional research output. We need an identifier that's attached to that, so it can be a DOI, a handle if it's in a repository, a PubMed ID, anything that is a unique identifier.

And then we look through the sources that we track – and we track thousands and thousands of sources. And as long as that's been mentioned with the identifier, it's very easy for us to pick it up. We do do some text mining so that artificial intelligence comes in to play where we'll look for an article title, but we do need the author as well, so we need two pieces of information to do that matching, and that gives us a positive match.

So that's a message for you if you're ever talking, you know, communicating, please use the DOI or use two pieces of information.

So this is what it looks like on the Taylor & Francis website, you can see the colourful donut and the colours within that donut represent the different types of sources. Sometimes people tend to get really focused on the score, but within your discipline you actually may have a relatively modest score, but be top in your area. So it's really important to go in and actually have a look

at those details and find out more.

I also wanted to just check with some of you because at the meeting on Monday, I talked to quite a few editors, and none of them had actually had a look at their Altmetrics for their journals. So there is a way of doing that. If you subscribe to an institutional product, you'll be able to get that. If you can't get access to the Altmetric database, you can always get in touch with your Taylor & Francis representative, and they'll help you to find out this. So you can do a quick search by publisher, or journal or collection, and that will bring back the results for you.

So if you've got the institutional product, you'll see a slightly different view where you can see a list. You see the institutional circle there, you can get a list of authors and departments at your institution. And so that's part of the institutional product, which is a fuller product.

So we were asked by a lot of our clients during the engagement and impact pilot, how Altmetrics could help them find out more. So when you're doing metrics for an impact or engagement study, you're perhaps not just looking for a citation count, what you're looking at is stories that can help you to craft a narrative. And the narrative is probably what makes up some of those interesting things behind. So a lot of universities would go and do things like the interview technique where they would talk to a researcher and say to them – what's interesting about your research? What has impact and who do you engage with? And they were very quick to say "my piece of research here has impact", but they couldn't provide evidence for that. Using the Altmetric database is a way that you can drill into that and you can find out where you've been mentioned in the news, who's talking about you, what are they saying. You should be also mindful that they could be saying something positive, or they could be saying something negative, so always good to check.

We are going to be doing more work on trying to analyse the text within those things, so that's coming in a future product.

Also what attention that matters to you? So you've got now some really influential bloggers, you know, they make the fact that somebody's blogged you really important. Or if you've published a piece of research that's got mentioned in a policy document, say in the World Health Organisation or one of those international sort of policy things that carries a lot of weight. So you may, you know, in this example here, may have a fairly modest Altmetrics score, but the fact you've been mentioned in 11 policy documents demonstrates that you've actually got a lot of impact. So it's those types of things that you can uncover using the Altmetrics database.

And this is the collaborative map where one person did a sort of interview with a researcher, and one of the things that they wanted him to demonstrate was his collaborative network. And so sometimes what you see is a lot of collaboration happens from previous institutions where you worked, but this actually shows you as an author who's actually picking you up. So we've got some different demographic maps there, we've got a Facebook, we've got a Twitter, we've got a news demographic map, and this is a map of the policy document. So you can see just,

you know, a nice visual representation, who's actually mentioned you and where those policy documents are from.

So to summarize, we make it easier for you to identify what might have impact, or to demonstrate engagement. We've got more content, we've got great visualisations, you can analyse the mentions by exporting them and looking for particular types of text. And we've got reach across a lot of sources. And also getting those results and doing analysis on them, we've got flexible reporting, we include a lot of visuals, you can save specific mentions and do benchmarking, look at competitor analysis, and you can share your narrative in a report, we've got a shareable report. Because we have a lot of philosophies around making data open, you can always extract the data using our API. So thank you.

James Hardcastle (Colwiz / Wizdom):

Right, so I'm James Hardcastle, and thank you for the introduction. So I used to work as part of Taylor & Francis. In fact, pretty much this time last year I gave a probably 50-minute presentation on metrics and recognition and incentives. I'm going to try and keep this one considerably shorter and talk a little bit more about Wizdom and some of the stuff that we're doing.

But before I start talking about products and the stuff that we're doing as part of Taylor & Francis, I just wanted to kind of help kind of set my scene in my own head as much as anything else, and think about the history of research, metrics and recognition, how people have been measured for their research over time. If you kind of think back to pre-Second World War times, it was peer impression that mattered, it was presentations, academic societies, and the general community of scientific research was small enough that everyone knew everyone.

The big explosion in scientific research happened in the '40s, '50s and '60s, you can't keep up, you can't know everyone, so you start to look for proxies of quality. People start to use journals, then number of journals increases beyond your ability to keep up with knowing which journal is good or bad, and Eugene Garfield created the impact factor in the 1960s I think, I'm getting a nod, that's positive. And people started to use those as proxies for journal quality.

And then we start to move into the kind of more modern age, we've got computers, we can quickly calculate metrics such as H indexes, we have citations ratios, field rated citation ratios for pretty much any dataset we want. And we can kind of start to look at social metrics and the wider old metrics, so policy documents etc.

But then we kind of come to things like the REF. I'm based in the UK and I talk to a lot of UK academics, and the REF gives them a large amount of paranoia and fear and drives a lot of their decision making. But we're moving back to this idea of peer assessment. People are reading impact statements, they're reading research, they're looking beyond the kind of metrics that we have initially. I just liked the kind of circle that we're almost moving back to, that idea of peer evaluation, peer assessment of research, rather than constantly looking for metrics as a

shorthand.

But perceptions matter almost as much as what people actually say. So we know in the UK, and I think Leon mentioned this earlier, that the funding councils in the UK have all signed DORA or will be signing DORA, which is the San Francisco Declaration on Research Assessment. Which basically says you shouldn't be using Impact Factor to evaluate work, you get highly cited articles in low Impact Factor journals, and zero cited articles in high Impact Factor journals. But pretty much every academic we speak to either is interested in Impact Factors and choosing where to publish, but they're also kind of using Impact Factors in things like tenure decisions, promotion committees all look for, or at least are perceived to be looking at the Impact Factors of peoples' publications, which venues have published their work, and using that as a proxy to decide effectively the quality of the research that someone is doing.

And this then leads to Tweets like this, so I stole this from EMBO Press, a researcher in Germany published what presumably is a very good paper, 40 citations a year, that's a really good level, but they're not allowed to include it as part of their research assessment exercise because the EMBO impact factor isn't high enough for them, and it's not high enough for their institutional threshold.

So we have this kind of mismatch between things like Impact Factors and how they're being applied, and actual kind of quality of work in terms of article level citations. And incentives matter as well, so I was kind of looking at this, and I was going to use an example on cycle helmets, but I couldn't find the research paper I wanted this morning. So instead I was going to talk briefly about the … and steal some slides from Euan, well a kind of concept that Euan Adie was talking about when he first founded Altmetric, things like the pressure to publish, the measure of how many articles you're producing, effectively is quite a perverse incentive because it encourages researchers to slice their work more thinly, and you end up with this move towards the minimal publishing unit. You want to publish five papers out of a set of research, when really you think you actually probably got three decent papers out of it. So how people set incentives can really change behaviour within the scientific community.

So we, Taylor & Francis, bought Wizdom or Colwiz back in the summer, and we are taking data on 60,000 odd journals, and I stole the leaflet because in Sydney I got the figures wrong and my memory is not good enough with jet lag to recount all of that. So 60,000 odd journals, 90 million publications, basically we're trying to take as much data as possible from everything. We've got about 1.5 trillion Australian dollars worth of grant funding into the database. We've got information on patents, clinical trials. And I'm less kind of prescient on remembering it's Valentine's Day today so I went for tennis as a theme for my slides because it was the Australian Open a few weeks ago. And you know, the cricket was a bit depressing for those of us in the UK.

So what we're doing is taking data across these sources and compiling it so we can see things like research output in tennis, which researchers, which universities are doing particularly well, who are the top most active researchers, what subject areas are people researching around

9

these topics, how has the grant location changed over time? Where are people publishing, is it all in the kind of biomechanics of tennis, there's stuff on culture, on sport and tourism and who is publishing this work as well?

We can also kind of tell you stuff about whether you can access again, stuff on institutions, the most recent publications, or we can look at how the subject field at kind of high level has changed over time. So we've moved from looking at an all time quite biological focus, into a slightly more social science focus in more recent, kind of last couple of years.

And we can also do this for a journal, so I wanted to try and pick a journal that wasn't going to kind of be one of anyone's in the room, so I went for PLoS Neglected Tropical Diseases. And we can start to do things like look at authors, who are new and returning, kind of understanding more about the dynamics within a particular journal, to help people understand trends within a set of data, understand who their authors are and which authors are doing particularly well or poorly. We can also look at citations, so I know going back to the REF, the Economics Research Council, the economics body to establish as part of the REF in 2014, did decide to use Impact Factor, because they said there was no better way to establish the predicted citations for an article on publication.

As a kind of complete tangent to this product, there was a company called Meta, which was acquired by the Chan Zuckerberg Foundation last year. And their big pitch to publishers was 'if you come and sign with us, we will tell you what the expected citation for this article is based on the author, the abstract, the title and the institution where they're published'. So tools like artificial intelligence and machine learning are allowing us to move away from those Impact Factors as the only measure of future citations to that kind of predicted bibliometrics for Leon's slide earlier, it does seem a bit odd, uncomfortable, to be evaluating people based solely on those kind of criteria. And again we can kind of keep looking at different types of datasets.

The one that I'm kind of most interested in I think is quite valuable data, and I think Dimensions and the interest in that also demonstrates this – is funding information. Who is getting funding, which topics are the funding bodies throughout the world allocating research grants to, who is receiving them and what type of research grants are they getting? Is it Intramural Grants, is it particular research projects, and how big are each of these individual grants?

So, I think in the session on Monday we had a brief conversation about looking at things … looking at grants and impacts based on grants, and if you were to look at citations or publication outputs per dollar of grant, you'll probably find that the arts and humanities do very well, because an arts and humanities grant is generally pretty small. In physics, life sciences, building something like CERN costs billions and billions of dollars, even though you are getting thousands of publications and thousands of citations out of it.

So there is a kind of a way of looking at this data and linking it together to give you an idea of impact in a slightly wider sense.