Universiteit Leiden
The Netherlands

The state of altmetrics: a tenth anniversary celebration

Christian, K.; Adie, E.; Derrick, G.; Didegah, F.; Groth, P.; Neylon, C.; ... ; Konkiel, S.

Citation

Christian, K., Adie, E., Derrick, G., Didegah, F., Groth, P., Neylon, C., ... Konkiel, S. (2020). *The state of altmetrics: a tenth anniversary celebration*. UK: DigitalCommons@University of Nebraska - Lincoln DigitalCommons@University of Nebraska - Lincoln. Retrieved from https://hdl.handle.net/1887/3277927

Version:	Publisher's Version
License:	Creative Commons CC BY 4.0 license
Downloaded from:	https://hdl.handle.net/1887/3277927

Note: To cite this publication please use the final published version (if applicable).

University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Copyright, Fair Use, Scholarly Communication, etc.

Libraries at University of Nebraska-Lincoln

10-2020

The State of Altmetrics: A Tenth Anniversary Celebration

Kathy Christian
Altmetric

Euan Adie
Altmetric

Gemma Derrick
University of Lancaster

Fereshteh Didegah
Karolinska Institute

Paul Groth
University of Amsterdam

See next page for additional authors

Follow this and additional works at: https://digitalcommons.unl.edu/scholcom

Part of the Intellectual Property Law Commons, Publishing Commons, Scholarly Communication Commons, Scholarly Publishing Commons, and the Social Media Commons

Christian, Kathy; Adie, Euan; Derrick, Gemma; Didegah, Fereshteh; Groth, Paul; Neylon, Cameron; Priem, Jason; Xu, Shenmeng; Zahedi, Zohreh; Theng, Yin-Leng; Hassan, Saeed-Ul; Aljohani, Naif R.; Bowen, Timothy D.; Patel, Vanesh M.; Haunschild, Robin; Bornmann, Lutz; Taylor, Mike; Ross, Liesa; and Konkiel, Stacy, "The State of Altmetrics: A Tenth Anniversary Celebration" (2020). *Copyright, Fair Use, Scholarly Communication, etc.*. 170.
https://digitalcommons.unl.edu/scholcom/170

This Article is brought to you for free and open access by the Libraries at University of Nebraska-Lincoln at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Copyright, Fair Use, Scholarly Communication, etc. by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Authors

Kathy Christian, Euan Adie, Gemma Derrick, Fereshteh Didegah, Paul Groth, Cameron Neylon, Jason Priem, Shenmeng Xu, Zohreh Zahedi, Yin-Leng Theng, Saeed-Ul Hassan, Naif R. Aljohani, Timothy D. Bowen, Vanesh M. Patel, Robin Haunschild, Lutz Bornmann, Mike Taylor, Liesa Ross, and Stacy Konkiel

This article is available at DigitalCommons@University of Nebraska - Lincoln: https://digitalcommons.unl.edu/scholcom/170

The state of altmetrics

A tenth anniversary celebration

October 2020

Altmetric

A message from Altmetric CEO Kathy Christian

Altmetric was founded in 2011 by my delightful predecessor, Euan Adie, for a simple purpose: to help researchers see the influence of their work in real time.

The pace of research may have accelerated over the past decade but this year it feels like we've kicked into a gear we didn't know we had. While 2020 has been frenetic on many fronts, it has been amazing to see how creativity and ingenuity can flow from a very difficult situation. It has come not just from researchers themselves, but also from those across the wider community: universities, governments, publishers, and businesses - all rapidly adapting the technologies and processes they use to manage research throughout its lifecycle.

The state of altmetrics discusses this rapid change and what we can expect in the future. We were lucky to receive contributions from global thought leaders. When I read them I was struck by the notion that the altmetrics of today are just the beginning. There are simply so many opportunities to serve and support the research community. Contributions touched on important, recurrent themes:

The need for responsible use of altmetrics in research evaluation. We cannot repeat the oversimplification of previous generations of research impact indicators;

The importance for altmetrics users, researchers, and providers alike to move beyond simple counting and correlations, towards contextualized indicators and qualitative impact evidence;

The field of altmetrics has flourished due to a diverse ecosystem of commercial and nonprofit innovators and data providers. We need to protect this ecosystem by ensuring altmetrics data remains open for researchers to study, and that data sources remain open for collection and integration into altmetrics services;

Data science techniques offer a great deal of promise for mapping the full landscape of engagement with research: how communities coalesce around particular topics, where altmetrics can be signals of deeper impact (e.g. disease monitoring), and identifying and dealing appropriately with "inorganic" sharing of research (e.g. bots); and

Altmetrics have grown tremendously in their use and recognition within academia, and there is likely much more growth to come.

Altmetric is proud to have spent the last decade as a change leader in the wider field of altmetrics. We have also been a student of the community, constantly learning how we can improve our products and services to better meet the needs of all.

In publishing *The state of altmetrics*, we hope to give back. In the following pages, you will hear from the authors of the original *Altmetrics manifesto*, researchers using altmetrics as social-spatial sensors to track emerging diseases, humanists keen to understand the ethics of altmetrics, and many others.

On behalf of the entire team at Altmetric, as you read this report I invite you to reflect upon an incredible decade, and imagine the future to come.

— Kathy Christian, CEO, Altmetric

Contents

1 Altmetrics, Ten Years Later **3**

Euan Adie (Altmetric (founder) & Overton)

2 Reflections on Altmetrics **5**

Gemma Derrick (University of Lancaster), Fereshteh Didegah (Karolinska Institutet & Simon Fraser University), Paul Groth (University of Amsterdam), Cameron Neylon (Curtin University), Jason Priem (Our Research), Shenmeng Xu (University of North Carolina at Chapel Hill), Zohreh Zahedi (Leiden University)

3 Worldwide Awareness and Use of Altmetrics **13**

Yin-Leng Theng (Nanyang Technological University)

4 Leveraging Machine Learning on Altmetrics Big Data **15**

Saeed-Ul Hassan (Information Technology University), Naif R. Aljohani (King Abdulaziz University), Timothy D. Bowman (Wayne State University)

5 Altmetrics as Social-Spatial Sensors **20**

Vanash M. Patel (West Hertfordshire Hospitals NHS Trust), Robin Haunschild (Max Planck Institute for Solid State Research), Lutz Bornmann (Administrative Headquarters of the Max Planck Society)

6 Altmetric's Fable of the Hare and the Tortoise **24**

Mike Taylor (Digital Science)

7 The Future of Altmetrics: A Community Vision **26**

Liesa Ross (Altmetric), Stacy Konkiel (Altmetric)

Altmetrics, Ten Years Later

Euan Adie
Altmetric (founder) & Overton, London, United Kingdom, e-mail: euan@overton.io

The year after I started working with altmetrics full time, I attended a fancy conference. At lunch I sat next to an academic from UCL. He asked what I did and then looked distinctly unimpressed. *Who really cares about that any of that stuff?* he wondered to the rest of the table. *Doesn't sound good for science.* He was here because of his current research, which he was clearly very passionate about. It was published in November, he explained excitedly to everybody, and already had 10,000 downloads, which was far more than anything else in the same journal…

Ten years on from the manifesto that gave altmetrics its name I'm pleased that this report is taking the time to look back and check: do people really care about any of this stuff? Is it good for research? And is it OK to be proud of your download counts, tactlessly at lunch or otherwise?

I think the answer to all three is yes. Of course researchers care about who is saying things about their work and about how that might compare to everybody else's: we're only human.

Funders, publishers and universities care too, for different reasons: what if people are saying bad things? Is it evidence we're reaching the right people? Can it help prove that we're achieving our mission?

To help with all this you need the underlying data – who's saying what – and, I'd argue, some quantitative indicators to help people benchmark themselves. Both fronts have seen a lot of effort over the past decade.

We've seen amazing work from Our-Research.org on ImpactStory, which presented a researcher controlled and orientated view of altmetrics data, and on projects like Depsy, which explored getting credit for people reusing your software. Plum Analytics cleverly started collecting citations from clinical information services like DynaMed, showing how you might target specific types of impact. Crossref and Datacite proposed and engineered a system that brought data auditability front and centre. At Altmetric we brought in university syllabi, patents and government policy as new sources.

There is no question that all of this new, underlying data has been good for individual researchers. When I was at Altmetric.com my favourite tweet was by an author engaging with a thread about her work that she'd found via its details page.

That isn't quite the kind of new metrics based discovery that the original manifesto talks about but it's still about connecting people and information in a useful way. For individual researchers this is what the most obvious value of altmetrics has ended up being: not solving any filter failure problems but presenting new data that wasn't previously available.

For larger groups and for discovery quantitative indicators are more useful and here work has moved slowly. That's not a bad thing: altmetrics is made up of dozens of heterogeneous data sources that cover a whole host of different intents, behaviours and potential outcomes. It's hard to build a comprehensive framework for understanding them, though through expert

"The rise of altmetrics has helped to legitimize people suggesting new measures for their work (or at least to push back against unsuitable ones)."

work from CWTS, the Scholarly Communications Lab, Max Planck, Indiana University, Warwick and elsewhere we're getting there.

What does the future hold? You'll find out in this report. I hope that as a motley mix of non-profits, companies and academic groups we'll continue to work together and learn from each other as a community rather than as straight up competitors. Personally I'm keen to see more tools and methods for summarizing and highlighting qualitative data - especially as it relates to the quality of the research being discussed, or can otherwise help public peer review - and for some kinds of data to be pooled as a common good, assuming we figure out who's willing to cover the costs of collecting and maintaining it in the long term.

In the meantime, and while recognizing that there's still so much more work to do, let's be proud of what we've achieved - the rise of altmetrics has helped to legitimize people suggesting new measures for their work, or at least to push back against unsuitable ones. The point of altmetrics data in research assessment was never to justify firing professors for not tweeting enough (this, maybe not unreasonably, was the #1 concern we'd hear from academics in the early days), but to give researchers working in fields poorly served by traditional citations something to work with.

Long may that continue.

Euan Adie founded Altmetric in 2011 and ran the company until 2017. He now runs Overton, which focuses on tracking policy related impact, and remains passionate about the way research is communicated online.

Reflections on Altmetrics

Gemma Derrick
University of Lancaster, Lancaster, England, e-mail: g.derrick@lancaster.ac.uk

Fereshteh Didegah
Karolinska Institutet & Simon Fraser University, Sweden / Canada, e-mail: fdidgah@gmail.com

Paul Groth
University of Amsterdam, Amsterdam, The Netherlands, e-mail: p.groth@uva.nl

Cameron Neylon
Curtin University, Perth, Australia, e-mail: cn@cameronneylon.net

Jason Priem
Our Research, Vancouver, Canada, e-mail: jason@ourresearch.org

Shenmeng Xu
University of North Carolina at Chapel Hill, Chapel Hill, NC, USA,
e-mail: shenmeng@email.unc.edu

Zohreh Zahedi
Leiden University, Leiden, The Netherlands, e-mail: z.zahedi.2@cwts.leidenuniv.nl

From the Editors:

In late 2010, a group of friends published an essay that quickly became a rallying cry for those disenfranchised with the scholarly status quo.

At just under 1000 words, the Altmetrics Manifesto <www.altmetrics.org/manifesto> recast previous theory-heavy musings on the "invocation of scholars on the Web"[1] into the idea of *altmetrics*: a mechanism for tracing how research is shared and discussed online.

The Manifesto envisioned a better way for research: a networked scholarly commons, representing the entirety of scholarship in all its diverse forms, leveraging the social web to help researchers navigate the deluge of articles (and presentations, figures, datasets, and so on) published each year. As almost an afterthought, the essay also suggested that one day altmetrics could be used to inform funding and promotion decisions.

The essay clearly struck a nerve. Since its publication in October 2010, the Manifesto has been cited a thousand times over, and spurred the development of an entirely new field of research, the creation of dozens of new companies, nonprofits, and labs. Perhaps most importantly, altmetrics have been adopted as an entirely new lens for academic evaluation and corporate decision-making.

[1] Cronin, B., Snyder, H.W., Rosenbaum, H., Martinson, A., & Callahan, E. (1998). Invoked on the Web. Journal of the American Society for Information Science, 49(14), 1319-1328.

A superficial reading of the Manifesto might suggest that the authors hold a Taylorist worldview. A cynical interpretation might suggest they think we should measure and map all aspects of research in order to rank scholars and promote ruthless efficiency alongside a "return on investment" mindset.

On the contrary, the original authors are responsible for some of the most interesting and provocative developments in "excellence" and "impact" studies to date, and have made an indelible mark on Open Science. Many researchers who have followed in their path are themselves mapping the boundaries of the unexpected interplay between science and society, and in some cases challenging academia's (incorrect) self-image of upholding an impartial meritocracy.

In the following pages, we celebrate the tenth anniversary of the publication of the Altmetrics Manifesto by interviewing three of the four original authors (Jason Priem, Paul Groth, and Cameron Neylon), along with current leading altmetrics researchers (Zohreh Zahedi, Gemma Derrick, Fereshteh Didegah, and Shenmeng Xu), to get their take on the current state and future of altmetrics.

Note: Some answers have been edited for length and clarity.

How has the use of altmetrics in practice differed from the original vision outlined in the Altmetrics Manifesto?

Paul: The original vision was very much focused on the need for better filters. In practice, altmetrics has primarily been seen as a way to get at measurements for other forms of impact or even just measuring faster forms of scholarly impact. Additionally, there's been a focus on traditional academic outputs (papers) rather than different kinds of outputs.

Cameron: This seems a shame because there is an incredible richness in the background information that social media and other digital traces of research use leave. Particularly in the current context, where on one side we see that the pandemic has both reduced the diversity of those engaging in research communication, and on the other that levels of exclusion of underrepresented groups from the academy has come into sharper focus, it would be good to see a broader commitment to a diverse use and application of this information.

All that said, the shift of social media-based indicators from a bizarre niche interest to a regular feature on university and publisher web pages perhaps shows that things can change.

"The shift of social media-based indicators from a bizarre niche interest to a regular feature on university and publisher web pages perhaps shows that things can change."

Zohreh: The manifesto was successful in extending the notion of impact beyond traditional measures by making the community aware of the limitations of current metrics. However, practical use of altmetrics for research evaluation and policy decisions is still finding its way. In particular, the potential of altmetrics in measuring societal impact of science and demonstrating the contribution that science has had to society at large is lacking.

Fereshteh: In practice, the aspects outlined in the Manifesto are well addressed so far, except for measuring the change science may make in different aspects of society in a feasible way. And I think it's mainly because qualitative analyses are lacking.

Jason: First, altmetrics have been hamstrung by a lack of open access to the event streams we want to analyze. "*Altmetrics*," reads the manifesto, use "*public APIs*." Those are pretty scarce these days. Most APIs are heavily restricted or very expensive to use at scale (e.g., Twitter) or they just don't exist: ever-greater swaths of the scholarly conversation now disappear into the profiteering maw of ResearchGate and other walled gardens, never to be seen again.

Second, the original vision was all about *filtering* the increasingly overwhelming din of scholarly communication. Sadly we are not even close to this. I don't think we really even have achieved the intermediate step *mapping* conversation's topology. Instead, we seem largely stuck on *counting* events. Counting was a necessary step. But I hoped that we'd be on to much more interesting, transformative uses by now.

Shenmeng: The manifesto discussed the use of altmetrics in a wide range of activities in scholarly communication. They all center around the idea of openness.

I'm viewing the Manifesto more as a call for researchers and practitioners in this community but less for users (e.g., like in the Leiden Manifesto). There have not been major deviations from the Manifesto's original vision. But there has been obstacles and resistance, particularly with openness. Without this really being resolved, it is difficult to talk about the ultimate utilities and best practices with this still young and rapidly evolving field.

The use and misuse of altmetrics, just as of any metrics, have always been an important issue of discussion. Although it was not a focal point in the manifesto, the warning message was clear. So far, so good. But we are not there yet.

Gemma: We need to address the wider social implications of adopting altmetrics for mainstream evaluative work. They were not considered in the original manifesto as at that time they were primarily promoted as an alternative to mainstream metrics. However, these issues are becoming increasingly important.

Mainstream metrics have had to address social issues such as responsible use and responsible evaluation. It would be reasonable, therefore, for altmetrics to be faced with the same questions around responsible and ethical use. These questions will also address questions around the seemingly lack of widespread adoption in evaluative practice, as was originally envisioned in the manifesto.

If you could change anything about the original Manifesto, what would you change?

Jason: This is a good question, but I tend to feel that what's done is done. I prefer to ask: what can we change about the present, and the future?

For me, the most important task right now is that we push for top-to-bottom openness in altmetrics and in scholarly communication in general. The fate of our species is hanging on the research enterprise. We cannot afford to balkanize and enclose it for a quick buck. We must fight that.

I do wish the manifesto hadn't been authored by four white guys (which was entirely my fault and not my coauthors').

Paul: In the original manifesto, I think we could have emphasized more the opportunity to use altmetrics to reflect diversity in scholarship and a bit less on the ability to capture interesting signals faster. I think involving a more diverse set of authors would have been good, too.

Shenmeng: Like Jason, I would probably elaborate a little more on the importance of openness in scholarly communication: open access, open data, open reviews, open communications, and open metrics. I would also very much like to see an updated Manifesto with more discussion about the benefits and drawbacks of altmetrics, and ideas on future directions for the field.

Fereshteh: I would add more about aspects of altmetrics that are under studied. Altmetrics could be used to study new (not necessarily scientific) interactions between science and society, how social networks coalesce around scientific outputs, and how these new networks affect making changes in the society.

Moreover, the interaction between science and society is mutual. Science can help society, and society can help science through instant reactions to scientific outputs, comments and annotations. The altmetrics Manifesto fully outlines the former, but it can also help with the latter.

Encouraging more context analysis would be another change I would like to make in the Manifesto. As a researcher, a question that I usually receive from the audience is: *What does a tweet about an article indicate?* There is a lot of "noise" and nonsense discussions around

"There is a lot of 'noise' and nonsense discussions around scientific articles, but we cannot ignore the effect of 'wisdom of the crowd'."

scientific articles, but we cannot ignore the effect of "wisdom of the crowd". Collective opinions about scientific outputs are not any less valuable than a systematic peer review.

Zohreh: I would add more about how different kinds of "actors" and users on social networks can a) be identified and b) studied to understand how they are interacting with and affecting each other. This helps in understanding the meaning and interpretation of what users' posts truly mean in practice.

Gemma: The development of flawed indicators to assess various interpretations of research 'excellence' is an affliction that has affected altmetrics as much as has mainstream metrics. I would have liked the original manifesto to have been able to anticipate this by addressing the potential misuse of altmetrics in the same regard as was considered mainstream metrics and, in particular, to how they might have exacerbated existing disparities in academia and research culture.

In addition, anticipating the influence of altmetrics when used as part of an evaluation narrative could have been addressed. Case study research[2] has shown how when altmetrics are used as part of a narrative to counter claims made by mainstream metrics, that the claims supported by mainstream metrics will be considered *more* reliable (not less). As a result, the inequalities in evaluation – although in some ways highlighted through altmetrics – will remain unresolved without human mitigation.

What do you think altmetrics as a field can learn from the discipline of data science, from both technical (e.g. graph analytics) and theory (e.g. ethics for big data) perspectives?

Shenmeng: On the one hand, similar to the first sentence in the Manifesto, no one can read every piece of news; no one can listen to every piece of music; no one can travel to every corner of the globe… Crowdsourced digital traces exist in many aspects of our daily life, and are widely used for filtering (e.g. recommendation) purposes to which machine learning is well-suited.

Also, despite the fact that the digital traces in scholarly communication have some unique features, machine learning based methods in social computing and data science could be applied throughout the data life cycle of altmetrics: data capture, database management, data cleaning and quality, as well as analytics and empirical research.

Zohreh: I agree. Data science can introduce advanced ways of data collection and data analysis and data mining for altmetrics. It offers new tools, methods, algorithms and capacity to deal with altmetrics "big data" and thus can help us better understand social media data, networks and nodes.

Fereshteh: I would love to see data science approaches taken that can help us find and track more publications and their altmetrics, improve data quality on all available platforms, and perhaps even find novel platforms like health forums for new data sources.

Paul: The data science field is frequently challenge or performance driven. E.g. "How did we perform on X & Y tasks?" This viewpoint is less prevalent within altmetrics research and use. But adopting a task oriented view might help us collectively to see altmetrics less as an end point (e.g., What is my paper's score?) and more as facilitating goals (e.g., Was my paper read by the people who would benefit from it the most?)

A second place to learn from data science would be openness about code and processing pipelines. We were lucky at the beginning of studying altmetrics and building altmetrics tools that commercial companies emerged to provide data for free. But in some sense that's meant fewer open source data and software pipelines.

[2] Aharony, Noa, et al. "Acceptance of altmetrics by LIS scholars: An exploratory study." *Journal of Librarianship and Information Science* 51.3 (2019): 843-851.

Jason: I also really love the trend in data science of open, replicable workflows, like you get with Jupyter notebooks[3]. It's easy to fork that analysis and then tinker with it yourself. I hope that more altmetrics researchers will continue to adopt this practice (as some have already).

One particularly promising part of the data science toolkit is machine learning. In fact I think machine learning (ML) isn't just *promising* for altmetrics—I think that the *point* of altmetrics should be to build ML-based models and tools. Machine learning techniques will eventually build a much-needed layer of abstraction between the researcher and the endless glassy sea of new literature.

Editor's note: See Chapter 4 for more on machine learning and altmetrics.

Cameron: Much of the more naive kinds of data science focus on categorization of static datasets or the search for correlations. I'm still frustrated that there hasn't been much work that gets beyond whether A correlates with B.

I remain convinced that we need to think differently about events that occur around research, thinking about them as *signals* of underlying processes. How do these signals propagate over time, and which channels (e.g., tweets vs usage vs citations) are coupled? How does that differ for disciplines, for output types, for languages? And beyond that, what does this tell us about enhancing the flow of information not so we get the most news coverage for a study but so that *the information gets to the right person at the right time*?

Ethics is also crucial. We need to think about consent and consequences of it.

Gemma: Yes, I agree. Altmetrics can certainly learn from data science, but altmetrics can also learn from the social sciences. Especially as the cultural integration of these tools – and what stops the further integration of these tools – into academic practice becomes more important. Only through a social science lens will we understand more about how tools are used, misused and what are the motivations behind this misuse.

In regards to ethics, for the social sciences in particular, there may be differences in the interpretation of an acceptable level of risk associated with using this data for individuals, than in the STEM disciplines. We saw this in our study of how ethics review committees and researchers assess risk in the use of social media in research[4]. This personal definition of risk is not always commensurate with the community's norms about responsible data use.

Where do you think altmetrics should be headed within the next five years, in terms of both use and technology?

Gemma: Agility is altmetrics' greatest advantage, and I would hope that there is increased innovation in creating new and refining existing altmetrics.

However, the use of altmetrics as data to underpin wider assessments of research quality/dissemination and excellence are, like mainstream metrics, blind to issues of equity and social justice queries during evaluation. Their use, unfortunately, still expands existing disparities in research production and academic culture than minimising their effects.

Shenmeng: I hope to see more researchers participate in open science/scholarship and build an open community together. For instance, register for an ORCID ID, publish replicable data, provide feedback on preprints, exchange ideas on social platforms, and communicate research with the public.

If we have an environment with better interoperability and transparency, we will be able to have more breadth and depth of altmetrics digital traces. Richer contextual clues will help altmetrics researchers better understand both what altmetrics can and cannot reflect.

[3] https://jupyter.org/

[4] Samuel, Gabrielle, Gemma E. Derrick, and Thed van Leeuwen. "The ethics ecosystem: Personal ethics, network governance and regulating actors governing the use of social media research data." *Minerva* 57.3 (2019): 317-343.

Fereshteh: On that point, I would like to see the study and use of altmetrics moving beyond counts, focusing on the rich content that social media can offer to the scientific community in order to build strong ties between science and society to help each other.

Zohreh: I can think of improving identification of social media actors (users) and their activities, contextualization of users' interactions, using opportunities of open science initiatives, improving access to rich data, and the development of advanced methods for dealing with big data analysis, among other things.

Paul: It would be nice to see an expansion of altmetrics to consider other forms of scholarly artifacts. For example, can we find good scientific videos that have contributed to a domain? Or can we understand what is the impact of all those EU project websites? Also, with more of our time spent interacting on video (e.g., virtual conferences), there's a real opportunity to create more real time altmetrics indicators that are up-to-date.

Jason: In the next five years, we have a big inflection point, for altmetrics and for scholarly communication in general. We need to choose what kind of *soil* we are going to put down, in which to grow the post-journal phase of scholarly communication.

We can have concrete parking lots with loads of little concrete planters for loads of sad little bushes, and each parking lot can have its own little concrete tollbooth. That's the way of closed-access, balkanized, gated data sources and data providers.

Or we can lay down one huge fertile field, where we throw every observation and data point from every experiment, and add in every bit of conversation or paper or poster or post we can find from everyone interested in this stuff, and we mix that all up together into a nutritious loam.

Now, that nutritious soil isn't very useful by itself. It's boring and dirty and kind of smelly. *But you can grow a forest in it.* So if we've got this soil laid down by 2025, then 2030 looks amazing. We can build an entire new scholarly communications system out and up from the raw substrate of interlinked data. We can create the tools we need to curate the increasingly uncurate-able mass of research, leveraging the vast flow of data to inform assistive filters.

Or we can let profiteers slice it all up for a quick buck. We've already seen the results of this in the larger social media sphere. It hasn't been pretty.

Where do you hope the field of altmetrics will be in the next 10-15 years?

Jason: In 15 years I think that altmetrics as such will gradually disappear, getting pulled into (and fueling) a larger revolution in research communication. The ideas in the manifesto will become real: we'll see real-time peer-review and curation based on automated distillation of expert conversation, moving beyond the paper[5] to create a vast and unowned decoupled journal[6].

This decoupled journal will chew up and digest all kinds of scholarly products, from tweets and conference talks to research notebooks and nanopublications to preprints and papers, to individual research observations and data points. It will output custom-curated streams and stories, each tuned to the interests of a particular researcher. Instead of reading 300 articles this year to keep up with my field, a machine will read a million for me–then summarize their findings, and generate novel hypotheses, perhaps even test them in robot labs[7]. Scientists will increasingly delegate normal science[8] to machines, leaving them space to apply higher-level creativity and inventiveness.

[5] Priem, J. (2013). Beyond the paper. *Nature* 495, 437–440. https://doi.org/10.1038/495437a

[6] Priem, J & Hemminger, B. (2012). Decoupling the scholarly journal. *Frontiers in Computational Neuroscience* 6, https://doi.org/10.3389/fncom.2012.00019

[7] Groth P, Cox J. (2017). Indicators for the use of robotic labs in basic biomedical research: a literature analysis. *PeerJ* 5:e3997 https://doi.org/10.7717/peerj.3997

[8] https://en.wikipedia.org/wiki/Normal_science

All of this will be based on interoperable open databases tracking billions of interlinked scholarly products, available to all, and powering hundreds of independent filtering services, both nonprofit and for-profit.

But to grow this ecosystem, we have to ensure that the soil is there. We must insist that all the products of scholarship are born open and stay that way. The rest, I think, will proceed quite smoothly.

Gemma: I would like altmetrics to focus on providing alternative narratives of excellence concentrating on the aspects of excellence overlooked by mainstream metrics. This can not be done by testing the validity of new altmetrics against mainstream metrics, but by creating alternative explanations and further discussions about what is research excellence. Only through the discussions that stem from a dashboard of metrics, illustrating where altmetrics and mainstream metric-based assessments differ, can discussions about competing notions of excellence and more robust and equitable evaluations be achieved. In a way, I would like altmetrics to play an active role in evaluators questioning their own judgements and biases, and to think as humans, not as machines.

"I would like altmetrics to play an active role in evaluators questioning their own judgements and biases,and to think as humans, not as machines."

Shenmeng: I hope to see the continuum of an increasingly open research environment. This will create unforeseen opportunities for infrastructural innovations—both technical and institutional ones.

Paul: It would be nice to see altmetrics as a core component within search and discovery systems for science.

Fereshteh: I hope as scientists, we'll be able to use altmetrics to better serve society in practice—using the data to detect public demands for the scientific community, i.e. their social needs and issues that need to be fulfilled by research.

Zohreh: I agree. I would love for altmetrics to be providing more useful insights that aid in the meaningful interpretation of data, so that they can be used to serve the community. I also think in the next 15 years, altmetrics will be indispensable in supporting policy makers and evaluators to make meaningful decisions.

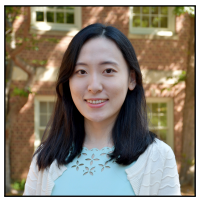
Gemma Derrick is a Senior Lecturer (Higher Education) at Lancaster University. She is Editor-in-Chief of *Publications* and Director of the Peer and Expert Review Laboratory. Her research focuses on the humanist angle of research culture, and the cultural adoption and orientation towards evaluation methods and tools.

Dr. Fereshteh Didegah is a researcher and data analyst currently working in the area of open science, societal impact of research, and research evaluation at Karolinska Institutet in Sweden and Simon Fraser University in Canada.

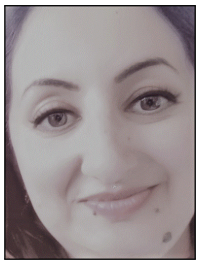
Paul Groth is a Professor of Data Science at the University of Amsterdam where he leads the Intelligent Data Engineering Lab.

Cameron Neylon is co-lead for the Curtin Open Knowledge Initiative which works on how to re-imagine what universities need to be in the 21st century and how we might encourage them to get there.

Jason Priem is co-founder of a nonprofit called Our Research, whose goal is to help accelerate the transition to Open Science. Our Research's latest project is Unsub, a data dashboard that helps libraries reassess and cancel their Big Deals, freeing up money for OA publishing.



Shenmeng Xu (PhD) is a Research Scientist at the School of Nursing at UNC Chapel-Hill. Shenmeng's research interests lie in the intersection of informetrics, social media, and scholarly communication. She is Chair of the ASIS&T Special Interest Group for the measurement of information production and use (SIG/MET).



Zohreh Zahedi holds a PhD in Social Media Metrics & Research Evaluation from CWTS, Leiden University. Her areas of expertise include social media metrics (altmetrics) and bibliometrics indicators, scientometrics, and research evaluation. Zohreh has published widely on the topic of social media metrics in leading peer-reviewed journals and she currently serves on the editorial board of the *Journal of Altmetrics*.

Worldwide Awareness and Use of Altmetrics

Yin-Leng Theng
Nanyang Technological University, Singapore, e-mail: tyltheng@ntu.edu.sg

Traditional metrics for scholarly and scientific publishing such as citation counts and impact factors have been criticised for being too shallow and too narrow. In order to go beyond limited comparisons of like-for-like and to become generally useful, we need better metrics, techniques, and tools to measure research outputs more accurately[1].

Worldwide, universities, research institutes, and policy makers realize that there is a need for scholarly metrics that allow equitable comparisons and that can be used to benchmark research. Efforts are needed to develop and identify more robust metrics for general use.

Today, with access to large volumes of data from diverse online platforms, it is possible to create alternative research impact metrics that capture domain-specific insights.

Altmetrics are an alternative, non-traditional class of metrics[2]. Altmetrics was first proposed in 2010 as a generalization of the concept of article-level metrics, and have been celebrated for their ability to trace the use of and engagement with scholarly articles, people, journals, books, data sets, presentations, videos, source code repositories, and web pages across both public and scholarly online spaces.

Though ten years in the making, altmetrics are still a new concept in some corners of academia. At the Centre of Healthy and Sustainable Cities (CHESS), Nanyang Technological University (NTU), we investigated the awareness and usage of altmetrics and the relationship between the usage of social media tools and the usage of altmetrics amongst academics worldwide.

In our study[3], we surveyed and received responses from 448 respondents including research staff (21%), professors (17%), students (13%), assistant professors (12%), and associate professors (10%) amongst others. Most participants were from Asia (45%), Europe (27%), and North America (14%).

Over three-quarters of researchers surveyed were familiar with traditional metrics like the Journal Impact Factor; slightly fewer (72%) were familiar with raw citation counts as an indicator of impact.

In comparison, the altmetric indicator with highest levels of familiarity was Mendeley readership (59%). Respondents were less familiar with altmetrics like "Number of 'likes' on social media" (30%) and "Number of mentions on blog posts" (19%).

Many respondents preferred traditional methods of promoting their research like publishing journal articles and presenting at conferences over sharing their work on social media.

"Researchers use altmetrics more when they are active social media users and when they perceive altmetrics to be useful."

[1] Sud, P., & Thelwall, M. (2014). Evaluating altmetrics. *Scientometrics*, 98(2), 1131–1143. https://www.doi.org/10.1007/s11192-013-1117-2

[2] Erdt, M., Nagarajan, A., Sin, S. C. J., Theng, Y. L. (2016). Altmetrics: An analysis of the state-of-the-art in measuring research impact on social media. *Scientometrics*, 109(2), 1117-1166. https://doi.org/10.1007/s11192-016-2077-0

[3] Aung, H. H., Zheng, H., Erdt, M., Aw, A. S., Sin, S. J. and Theng, Y. L. (2019), Investigating familiarity and usage of traditional metrics and altmetrics. *Journal of the Association for Information Science and Technology*. https://doi.org/10.1002/asi.24162

Importantly, we learned that researchers use altmetrics more when they are active social media users and when they perceive altmetrics to be useful. For instance, scholars may better appreciate the benefits of altmetrics when they use altmetrics to identify high quality research outputs that are relevant to their own research areas, or when altmetrics are included in evaluation guidelines. We suggest that new and systematic evaluation criteria that includes altmetrics should be established to increase the perceived benefits of altmetrics among scholars, and thus increase their use.

Altmetrics offer promise over traditional methods of measuring research impact that can be biased against certain disciplines. However, altmetrics has its own limitations. The quality of altmetrics needs to be enhanced (i.e., "the lack of accuracy, consistency and replicability of various altmetrics, which is largely affected by the dynamic nature of social media events"[4]), and appropriate strategies and efforts should be made by research institutions and publishers to improve the familiarity and usage of altmetrics in academia, if there is hope that altmetrics be widely adopted alongside traditional metrics.

Acknowledgements Our research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its Science of Research, Innovation and Enterprise programme (SRIE Award No. NRF2014-NRF-SRIE001-019).

Dr. Theng Yin Leng is Associate Provost (Faculty Affairs) and Professor of Information Studies at the Wee Kim Wee School of Communication and Information, Nanyang Technological University (NTU, Singapore). Her main research interest is experimental Human-Computer Interaction in understanding users and their interactions of interactive systems/devices.

[4]Haustein, S. Grand challenges in altmetrics: heterogeneity, data quality and dependencies. *Scientometrics* 108, 413–423 (2016). https://doi.org/10.1007/s11192-016-1910-9

Leveraging Machine Learning on Altmetrics Big Data

Saeed-Ul Hassan
Information Technology University, Lahore, Pakistan, e-mail: saeed-ul-hassan@itu.edu.pk

Naif R. Aljohani
King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia, e-mail: nraljohani@kau.edu.sa

Timothy D. Bowman
Wayne State University, Detroit, MI, United States of America,
e-mail: timothy.d.bowman@wayne.edu

Nearly a decade ago, the term altmetrics was coined as an alternative form of representing online social engagements with scholarly documents to filter the vast amount of information being disseminated in social media and other contexts for the timely measurement of a document's (in a broad sense) impact[1]. Since the inception of the term, altmetric scholars from a variety of disciplines continue to develop, test, and theorize new means of collecting, analyzing, and measuring online activities focused on the consumption and dissemination of scholarly documents in a variety of online contexts[2] – with over 160 authors publishing nearly a thousand altmetrics publications as of June 2020.[3]

Given that altmetrics data satisfies the three main characteristics of big data, i.e., volume (data size), velocity (rate of data increase), and variety (data from different social media platforms), it is surprising that only seven altmetrics publications to date address one of the most popular big data topics of our time: machine learning. We believe there is a lot of opportunity to use new tools and techniques offered by the advancements in machine learning and big data platforms in the search for meaning in altmetrics data, especially by shifting focus from counting mentions of research within particular altmetrics sources to studying who is discussing research online.

What is machine learning?

Machine learning is a subset of the field of data science that uses algorithms to make predictions based on existing data. Some techniques like logistic regression are already widely applied in quantitative research fields; other techniques like k-means clustering are less often used outside of computer science and engineering research.

Though it's beyond the scope of this essay to give a full explanation of the many machine learning approaches in use today, there are three areas of machine learning that we believe can be useful in altmetrics studies:

[1] Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010). Altmetrics: A manifesto. Available at: http://altmetrics.org/manifesto/

[2] Sugimoto, C. R., Work, S., Larivière, V., & Haustein, S. (2017). Scholarly use of social media and altmetrics: A review of the literature. Journal of the Association for Information Science and Technology, 68(9), 2037-2062.

[3] Against the search term "altmetric" or "altmetrics" when matched with titles, abstracts, or keywords of all the publications indexed by Scopus.

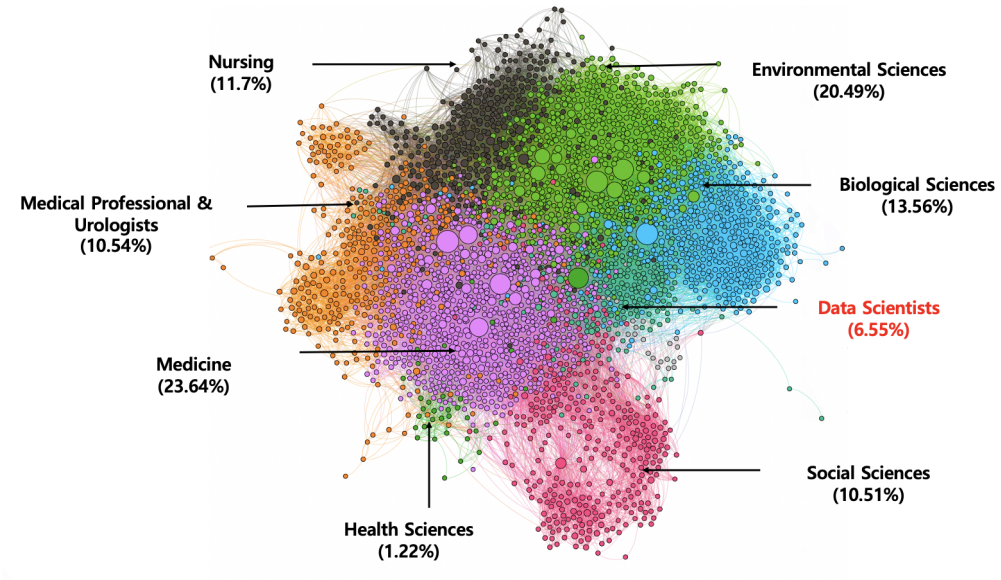


Figure 4.1: Visualization of the 'mention' network of 1.4 million tweets across major communities in altmetrics. For more information, read Said et al. 2019.

- **Cluster analysis**, in which an algorithm selects similar groups of data points based upon a number of options like their connectedness or distance to other members of the group, or their distance from a central point in the group; and
- **Feature learning**, in which an algorithm identifies the properties (features) of a dataset that have the most predictive power—for example, the shape of a leaf is a feature that allows a user to identify a tree.
- **Deep learning**, an approach that repeatedly feeds data through an algorithm to identify high-level features in a dataset that aren't readily apparent. Deep learning can thus be used for certain tasks that humans tend to be good at, like image identification.

Here, we share a few recent studies that highlight the promise of these machine learning approaches in pushing the boundaries of altmetrics research.

Using cluster analysis to identify user communities sharing research

Our recent work[4] - which is an extensive study of a Twitter mention network that discussed research published in 2015 - uses the Louvain method for community detection[5] to show the diverse communities within and outside academia who engage with scholarly documents on Twitter. Surprisingly, the most talked-about workers of the 21st century, data scientists, make up only a small proportion of those discussing research online, with just over 6% coverage.

"There is value in analyzing altmetrics data through new techniques in order to shed light on unusual phenomena."

The altmetric mention network (Figure 4.1) contains over 1.4 million tweets, with each node representing a Twitter user account, with the colors representing a specific user community and with the size of each node representing the user's influence on the network. The most influential users in this tweet altmetric network are from medicine, environmental science, and biological sciences, and none of the influential Twitter users are data scientists. The top fifteen influential Twitter users in this study were unexpectedly either associated with a publisher or an organization, rather than individuals.

[4] Said, A., Bowman, T. D., Abbasi, R. A., Aljohani, N. R., Hassan, S. U., & Nawaz, R. (2019). Mining network-level properties of Twitter altmetrics data. Scientometrics, 120(1), 217-235.

[5] Vincent D Blondel et al J. Stat. Mech. (2008) P10008 https://iopscience.iop.org/article/10.1088/1742-5468/2008/10/P10008

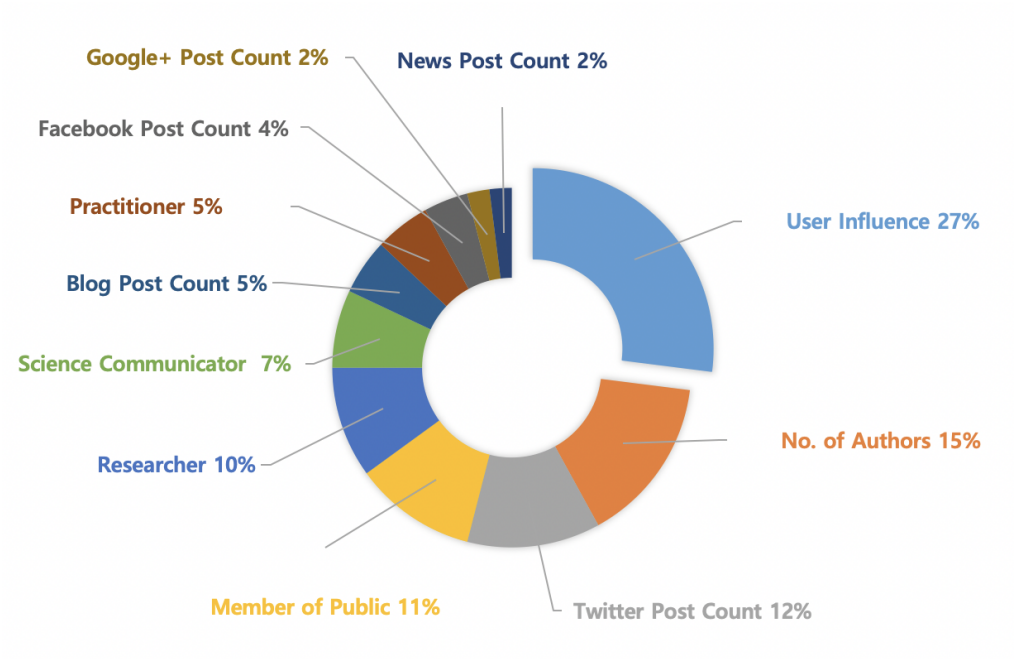


Figure 4.2: Importance for each feature to predict highly cited paper using Machine Learning Model i.e. Extra-Tree Classifier

This discovery was surprising, given that data scientists are an active Twitter community who regularly discuss technical concepts based on research. Moreover, few jobs dominate the news as often as data scientists, so it is unexpected to see this much-discussed community was unranked among the most influential Twitter users discussing research. Clearly, there is value in analyzing altmetrics data through new techniques in order to shed light on unusual phenomena.

Feature learning can predict publications with high citation rates

Another recent study found that user influence within an altmetric Twitter mention network is the most crucial feature (Figure 4.2) of an altmetric dataset that can predict publications with high citation rates. In the study, user influence is measured as "those [users] which, when deleted, cause the greatest drop in the eigenvalue of the graph". The authors argue that the effect of influential users may give rise to bias in generating social usage data and future citation counts. A more comprehensive understanding is needed to investigate other factors that can directly or indirectly influence Twitter altmetric counts - which alone covers over 90% of the total altmetric mentions.[6]

Deep learning to identify bots

Deep learning can be utilized to better understand the underlying social media networks associated with altmetrics, for example, to identify bots. One recent study[7] did just that, using a type of deep learning called Graph Convolutional Neural Networks (GCNN) to learn the structure of a social media network, using the edges among the nodes instead of any semantic

[6] Hassan, S. U., Imran, M., Gillani, U., Aljohani, N. R., Bowman, T. D., & Didegah, F. (2017). Measuring social media activity of scientific literature: an exhaustive comparison of Scopus and novel altmetrics big data. Scientometrics, 113(2), 1037-1057.

[7] Aljohani, N. R., Fayoumi, A., & Hassan, S. U. (2020). Bot prediction on social networks of Twitter in altmetrics using deep graph convolutional networks. Soft Computing, https://doi.org/10.1007/s00500-020-04689-y

attached to the nodes (e.g. language of a mention[8]). By mapping the structure using GCNN, the authors were able to predict bots that affected 87% of tweets from a Twitter network's dominant communities, with up to 71% accuracy.

However, deep learning is not a panacea. To avoid the pitfalls of other domains that have adopted machine learning models as a black box, researchers need to scrutinize the underlying social media networks by selecting the right model with an in-depth understanding of features.

The future: Better understanding users' effect on altmetrics using machine learning techniques

Given the promise in using machine learning for altmetrics research, it's time to shift altmetric researchers' focus from the research object to users (persons, groups, or organizations) and their activities. Using machine learning techniques, altmetric researchers can better illuminate the "social" aspect of social networks and user activities, focusing their lens of analysis on who consumes and disseminates scientific documents and how else they use research in these contexts.

By shifting the focus of altmetric research, researchers can provide a broader understanding of why users consume and disseminate scholarly documents, to whom they share scholarly documents, and why users of these online platforms undertake these activities. This will lead to a better understanding of the types of impact these scholarly objects can have both within and outside of academia.

Altmetric data allows us to look at the interaction with a scientific object by examining how users access, apply, and appraise research.[9] Machine learning approaches to altmetrics analysis can further explore this idea by ingesting additional data that allows for the analysis of a user's social network, their self-proclaimed profile description, and their other activities within the online platform(s). Extending the dataset allows for the expansion of analysis from simply examining the activities surrounding the scholarly output to also examining the activities of the users themselves.[10]

We believe that machine learning and deep learning can be leveraged to provide a holistic view of scholarly communication by providing contextual representation of user activities on different social media platforms using the models analogous to word embeddings, developed by Mikolov et al. (2013) at Google,[11] which is capable of capturing the context of a data object, semantic and syntactic similarity, and relation with other data objects in the information space.

There is much promise in exploring new analytical techniques like machine learning to expand what we can learn about altmetrics. We urge the community to take a closer look at how machine learning and deep learning might move us toward developing more accurate predictive analytics rooted in altmetrics. We also caution that these exciting new technologies should be

[8] Hassan, S. U., Saleem, A., Soroya, S. H., Safder, I., Iqbal, S., Jamil, S., Bukhari, F., Aljohani, N. R., & Nawaz, R. (2020). Sentiment analysis of tweets through Altmetrics: A machine learning approach. Journal of Information Science, https://doi.org/10.1177/0165551520930917

[9] Haustein, S., Bowman, T.D., Costas, R. (2016). Interpreting "altmetrics": Viewing acts on social media through the lens of citation and social theories. In Cassidy R. Sugimoto (Ed.). Theories of Informetrics and Scholarly Communication. Berlin, Boston: De Gruyter Mouton

[10] Díaz-Faes, A.A., Bowman, T.D., & Costas, R. (2019). Towards a second generation of 'social media metrics': Characterizing the interactions of Twitter communities of attention around science. PLOS ONE, 14(5): e0216408. https://doi.org/10.1371/journal.pone.0216408

[11] Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In Advances in neural information processing systems (pp. 3111-3119).

carefully monitored to prevent misuse and abuse similar to what has been widely reported in recent months from other sectors.[12] [13] [14]

Dr. Saeed Ul Hassan is the Director of Scientometrics Lab and Associate Professor at Information Technology University, with more than 15 years of hands-on experience in machine learning, deep learning, and software development.

Dr. Naif Aljohani is an Assistant Professor in the Information Systems Department at King Abdulaziz University in Saudi Arabia. His research interests include social networks, data mining, and linked open data.

Dr. Timothy D. Bowman is an Assistant Professor in the School of Information Sciences at Wayne State University. Tim has written and presented widely about scholarly communication, bibliometrics, and altmetrics.

[12] Tzachor, A., Whittlestone, J., Sundaram, L. et al. Artificial intelligence in a crisis needs ethics with urgency. Nat Mach Intell (2020). https://doi.org/10.1038/s42256-020-0195-0

[13] Cheatham, B., Javanmardian, K. & Samandari, H. (2019). Confronting AI risks. McKinsey Company. https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/confronting-the-risks-of-artificial-intelligence

[14] Reiland, R. (2018). Artificial Intelligence Is Now Used to Predict Crime. But Is It Biased? Smithsonian Magazine. https://www.smithsonianmag.com/innovation/artificial-intelligence-is-now-used-predict-crime-is-it-biased-180968337/

Altmetrics as Social-Spatial Sensors

Vanash M. Patel
West Hertfordshire Hospitals NHS Trust, London, United Kingdom,
e-mail: vanash.patel06@imperial.ac.uk

Robin Haunschild
Max Planck Institute for Solid State Research, Stuttgart, Germany,
e-mail: R.Haunschild@fkf.mpg.de

Lutz Bornmann
Administrative Headquarters of the Max Planck Society, Munich, Germany,
e-mail: lutz.bornmann@gv.mpg.de

Citations have long been used to measure research impact, but they have been criticized by many for not being able to reflect the broader impacts of research, such as educational, cultural, environmental, and economic impact[1] [2]. The exponential growth of digital technology, particularly in the last decade or so, has enhanced scientific productivity, significantly expanding the methods in which research is communicated, archived, and assessed.[3]

One such example is that researchers increasingly use Twitter as a communication platform, and tweets can contain mentions of scientific papers[4]. Twitter mentions—defined as "direct or indirect links from a tweet to [research outputs] online"[5] [6] —can become part of a rapid dialogue between users which may express and transmit academic impact and support traditional citation analysis. Twitter mentions can reflect a broader discussion of research that crosses traditional disciplinary boundaries and represents "attention, popularity or visibility"[7] rather than what is commonly understood to be scholarly influence.

Studies have shown that tweets can be used as 'social sensors', which is the concept of replacing a physical sensor in the real world with social media analysis. Tweets can be regarded as sensory information and Twitter users as sensors. Studies have demonstrated that tweets analysed as social sensors can provide insight into major social and physical events like

[1] Bornmann, L. (2016). Scientific revolution in scientometrics: The broadening of impact from citation to societal. In C. R. Sugimoto (Ed.), Theories of informetrics and scholarly communication (pp. 347-359). Berlin, Germany: De Gruyter.

[2] Bornmann, L., & Haunschild, R. (2017). Does evaluative scientometrics lose its main focus on scientific quality by the new orientation towards societal impact? Scientometrics, 110(2), 937-943. doi:10.1007/s11192-016-2200-2

[3] Markusova, V., Libkind, A., & Bogorov, V. (2018). Usage metrics vs classical metrics: analysis of Russia's research output. Scientometrics, 114(2), 593-603.

[4] Priem, J., & Costello, K. L. (2010). How and why scholars cite on Twitter. Proceedings of the American Society for Information Science and Technology, 47(1), 1-4. doi:10.1002/meet.14504701201

[5] Priem, J., & Costello, K. L. (2010). How and why scholars cite on Twitter. Proceedings of the American Society for Information Science and Technology, 47(1), 1-4. doi:10.1002/meet.14504701201

[6] Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). Earthquake shakes Twitter users: real-time event detection by social sensors. Paper presented at the Proceedings of the 19th international conference on World wide web.

[7] Mas-Bleda, A., & Thelwall, M. (2016). Can alternative indicators overcome language biases in citation counts? A comparison of Spanish and UK research. Scientometrics, 109(3), 2007-2030. doi:10.1007/s11192-016-2118-8

earthquakes[8], sporting events[9], celebrity deaths[10], and presidential elections[11]. Twitter data can contain location information that can be converted into geo-coordinates and be spatially mapped. In this way, tweets can also be used as social-spatial sensors to demonstrate how research diffuses within a population[12].

We have studied how Twitter data mentioning research can be used as social-spatial sensors in several scenarios. We set out to investigate whether research on certain diseases reaches the populations that are especially affected by the diseases. The diseases we included in our study were tuberculosis, human immunodeficiency virus (HIV), and malaria, which the World Health Organisation (WHO) ranks as the top three causes of death worldwide by a single infectious agent. Whilst we were undertaking our study, a novel infectious agent called severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) emerged causing Coronavirus disease 2019 (COVID-19), and the WHO declared this disease a pandemic on 11 March 2020[13]. The global impact of COVID-19 compelled us to also include this disease in our research[14].

"Higher Twitter activity was associated with higher incidence rates of the diseases and number of publications."

We used three kinds of data in our study:

- Incidence rates of the diseases (WHO)
- Publications related to these diseases (bibliographic databases)
- Tweets that mention these publications (Altmetric.com API & Twitter API)

Plotting user locations for tweets that mentioned publications related to the diseases on worldwide maps revealed that high Twitter activity corresponded to high incidence rates of the diseases (Figure 5.1). Additional statistical analysis significantly showed that higher Twitter activity was associated with higher incidence rates of the diseases and number of publications, although the influence of the number of publications was greater than the incidence rates of the diseases.

Thus, we believe that Twitter mentions of research can be used as social-spatial sensors to monitor research diffusion for diseases that have significant worldwide impact. Our results suggest that research broadcast through Twitter reaches populations that are concerned about the diseases in this study.

Our study can serve as proof-of-concept for future studies on Twitter data as indicators of research impact[15]. We welcome the community to test our approach, and suggest that future studies could incorporate thematic analysis of the tweets in terms of their content (e.g., are tweets referring to disease testing, therapies, or vaccines), quality (e.g., whether tweets are referring to original research or correspondence), or who tweeted these (e.g., researchers, members of the public, universities, or pharmaceutical industries).

Our study has illustrated a glimpse of the relationship between diseases, research outputs, and Twitter activity, and conveys the importance of social media platforms for knowledge transfer. Therefore, we urge governments to pause censorship of social media platforms such as Twitter

[8] Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). Earthquake shakes Twitter users: real-time event detection by social sensors. Paper presented at the Proceedings of the 19th international conference on World wide web.

[9] Takeichi, Y., Sasahara, K., Suzuki, R., & Arita, T. (2014). Twitter as Social Sensor: Dynamics and Structure in Major Sporting Events. Paper presented at the ALIFE14.

[10] J. Sankaranarayanan, J., Samet, H., Teitler, B. E., Lieberman, M. D., & Sperling, J. (2009). TwitterStand: news in tweets. Paper presented at the Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems.

[11] Shamma, D. A., Kennedy, L., & Churchill, E. F. (2009). Tweet the debates: understanding community annotation of uncollected sources. Paper presented at the Proceedings of the first SIGMM workshop on Social media.

[12] Bornmann, L., Haunschild, R., & Patel, V. M. (2020). Are papers addressing certain diseases perceived where these diseases are prevalent? The proposal of Twitter data to be used as social-spatial sensors. http://arxiv.org/abs/2004.13974.

[13] Ghebreyesus, T. A. (2020). WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020 [Press release]. Retrieved from https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020

[14] Patel, V., Haunschild, R., Bornmann, L., & Garas, G. (2020). A call for governments to pause Twitter censorship: a cross-sectional study using Twitter data as social-spatial sensors of COVID-19/SARS-CoV-2 research diffusion. medRxiv, 2020.2005.2027.20114983. doi:10.1101/2020.05.27.20114983

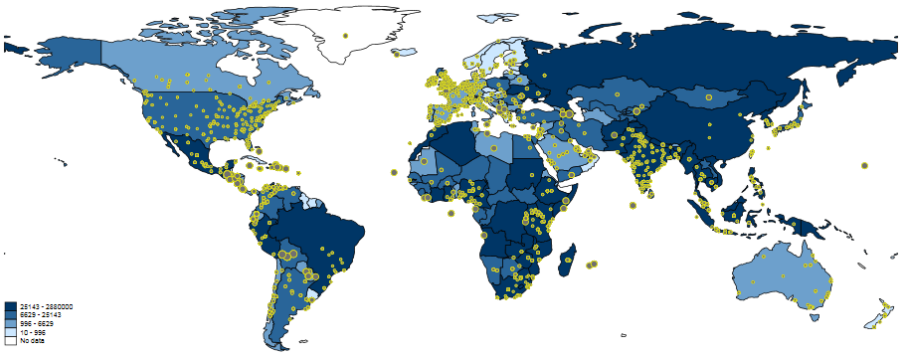
[15] Thelwall, M. (2017). Web indicators for research evaluation: A practical guide. London, UK: Morgan & Claypool.

during these unprecedented times to support the scientific community's battle against diseases such as COVID-19, using monitoring and prediction methods like altmetrics.

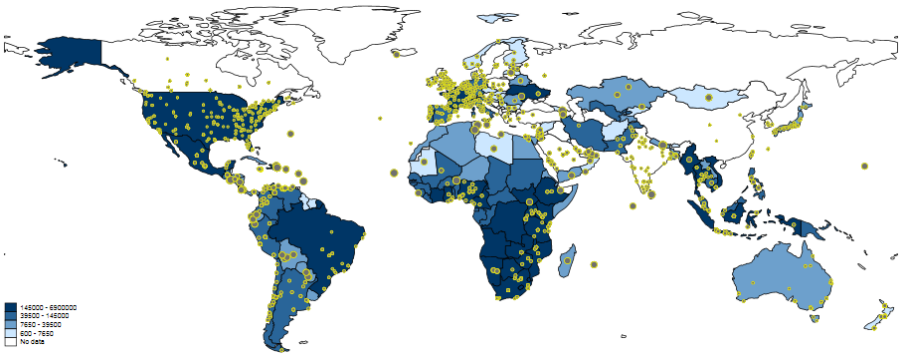
Vanash Patel is a consultant colorectal surgeon at West Hertfordshire Hospitals NHS Trust, UK. During his surgical training, he completed a PhD at Imperial College London, investigating academic performance metrics. Vanash is currently an honorary research fellow at Imperial.

Robin Haunschild is a chemist by training and joined the Max Planck Society in 2014. His current research interests include the study of bibliometrics and altmetrics as well as their application to specific fields of natural sciences, e.g., chemistry and climate change.

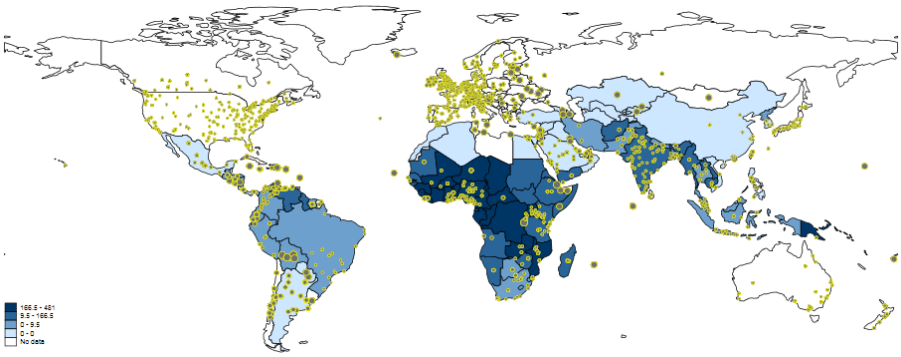
Lutz Bornmann is a habilitated sociologist of science and works at the Division for Science and Innovation Studies of the Max Planck Society. His current research interests include research evaluation, peer review, bibliometrics, and altmetrics.



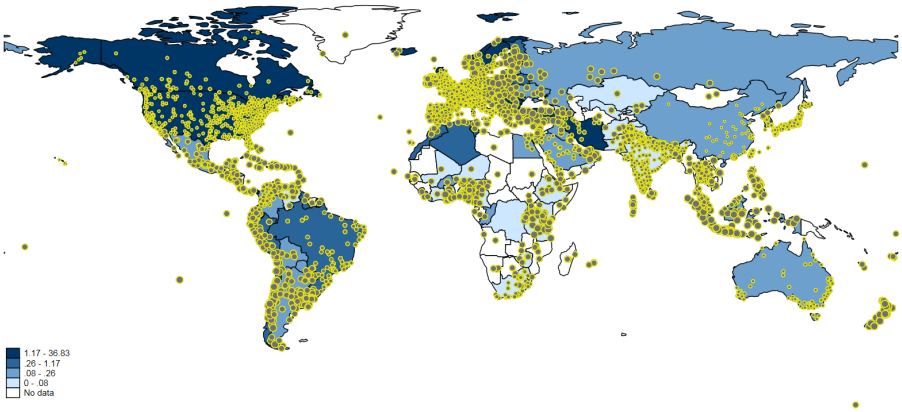
(a) Tuberculosis



(b) HIV



(c) Malaria



(d) COVID-19

Figure 5.1: Tweeting on publications dealing with (A) tuberculosis, (B) HIV, (C) malaria and (D) COVID-19 worldwide. Each tweet is inversely weighted with the number of publications published by authors in the corresponding country: the larger the dots, the smaller the research activity. The countries are colored according to the incidence rates of the disease. For some countries, e.g. Greenland, no data are available. Countries such as China and Iran block internet access to Twitter or its content.

Altmetric's Fable of the Hare and the Tortoise

Mike Taylor
Digital Science, London, United Kingdom, e-mail: m.taylor@digital-science.com

Once upon a time in a field not very far away, there lived a hare and a tortoise. There were many hares, packed with all kinds of ideas, and they raced around the countryside making lots of twittering sounds and crashing through hedges.

In the same field, there lived some tortoises. There weren't as many of them, because it takes a long time for a tortoise to grow up, and not all fields are kind to tortoises. But the tortoises are steady, thoughtful animals, and their companies in the field valued their sagacity and insights.

One day, some visitors came to the field and decided to have a race. They would put the hares against the tortoises, and see who won. The first race was just from the hedge to the nearby apple tree, and the hares won easily. The second race was from one side of the field to the other, and although the tortoises at least got off the starting line, the hares won that one too.

There was much debate. It didn't seem fair to match these two groups of animals against each other, until someone suggested that since there were lots of hares, and they were valued for their speed and noise; but that tortoises were valued for their wisdom and knowledge of the field.

The new day dawned brightly, and a new race was staged. This time, instead of sprinting across the field, the animals would make their way around the hedge, and would carry on racing until lunchtime.

The starting flag was waved: and the hares sprinted off! Around the field they ran, lapping the tortoises with ease. But on the second lap, one of the hares got bored, and started nibbling grass. The tortoises overtook it. On the third lap, one of the hares crashed through a hedge, and ran away. The tortoises made their way past the hole, without even looking. On their third lap, two of the hares started fighting, and the tortoises watched for a moment, blinked, and carried on.

On its fourth lap, the final hare ran out of breath, and seeing that the tortoises were still far away, decided to rest. The tortoises crept up, not daring to breathe, and made their way back to the starting line - just in time for a delicious lunch of lettuce.

The analogy between hares and journal articles, and tortoises and books might seem a little far-fetched, but as altmetrics marks its tenth anniversary, it seems a timely moment to reflect on how books are seen through the lens of altmetrics.

As I and my co-authors of *The State of Open Monographs*[1] explored in 2019, there are a number of challenges faced by people trying to assess the role of the academic book in the scholarly environment. The metadata that enables books to be tracked and reported is both

[1] Grimme, S., Taylor, M., Elliott, M. A., Holland, C., Potter, P., Watkinson, C.. (2019). The State of Open Monographs. Digital Science. https://doi.org/10.6084/m9.figshare.8197625.v4

more complex than for journal articles - and also less well-populated. Books are fewer in number, and whereas historically, sales figures dominated publishers' analyses, the move towards increasing Open Access publishing in books is depriving people of this otherwise important metric. All importantly, books have an impact profile very different to that of the journal article. The latter - the hare in our fable - gets its initial impact fast, but in many fields has ceased acquiring both citation data and altmetric data as early as two years after publication. Often our tortoises - our books - have barely started making an impact aged two.

"Books have an impact profile very different to that of the journal article."

The many difficulties in acquiring book data has meant that - just like our tortoise - research about the altmetrics of books has been slow off the mark. Fortunately, both the books and scientometrics communities are increasingly understanding the value in analysing this data; and the movement towards OA publishing has led to increased work to understand the role of the scholarly book.

This enthusiasm, together with the innovations made by Altmetric, Dimensions and others has led to an increase in research in this area. We know, for example, that Open Access books show a significant advantage over non-Open Access books across a broad range of altmetric indicators. Research shows that book impact continues to grow over many years, with the implication that this rate of growth may outstrip equivalent research articles over a period of ten years or more. Books appear to show up preferentially in policy sources, and the value of innovative books-only data sources has only started to become apparent.

Ask any "book person", and they will jump to the defense of the book, and argue for the unique role played by books in many fields, not just the arts, the humanities and the social sciences (as an ardent book person myself, my first projects were in statistical mechanics and the chemistry of coal). But we live in a time when all forms of scholarship deserve to be assessed, and both quantitative and qualitative approaches are now commonplace. Whether we book people like it or not, we cannot take the importance of books 'as read'.

The increased focus on the importance of books, the innovations in data collection and analysis, as well as the acknowledgement that books and their impact are distinct, and require a distinctive analysis mean that the years to come will be fruitful for the understanding of book impact.

Mike Taylor is head of metrics development at Digital Science, where he specializes in quantitative and qualitative analyses of academic trends.

The Future of Altmetrics: A Community Vision

Liesa Ross
Altmetric, Baltimore, MD, USA, e-mail: liesa@altmetric.com

Stacy Konkiel
Altmetric, Minneapolis, MN, USA, e-mail: stacy@altmetric.com

A lot of amazing innovation in altmetrics comes from our forward-thinking community. We recently attempted to tap into this group wisdom by surveying 410 individuals to identify a collective vision for the future of altmetrics.

In our survey, we asked questions around topics like: Do you expect your use of altmetrics to change soon? What about changes in the use of altmetrics overall in the community? Will altmetrics see a reinvention, or continue to be "more of the same"? We were surprised and delighted by many of the other predictions offered.

In this essay, we share survey highlights that explain the current state of altmetrics and where we collectively expect to end up.

Currently: The community is largely academic and "Global North"

We ran two versions of a survey, one for customers and one for the larger community (non-customers). Overall, survey respondents in both groups hailed from throughout the world (with just more than 60% living in North America and Europe), occupying a variety of technical and non-technical positions at universities, publishers, and corporations. Most respondents have been using altmetrics in some form for at least 2 years. A majority of survey respondents work in academic institutions (> 50%), and are either researchers or librarians (> 60%). Keep this skew in mind as you interpret the responses below.

Currently: Altmetrics are used to track and explain research impact

Customers and non-customers alike tend to use altmetrics for the following purposes (listed in order from most to least common use): Track research impact, Help explain research impact to others, Discover research trends, Identify influencers, Competitive analysis, Plan marketing and communications activities, and Make decisions.

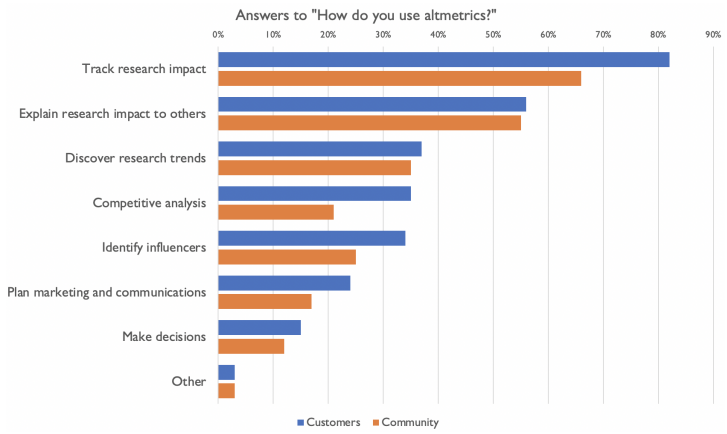


Figure 7.1: Tracking and explaining research impact are the most popular uses of altmetrics.

Half of Altmetric customers are using altmetrics just as often now as they had in the past, and one-third of customers are actually using altmetrics more often. Among the broader community, roughly as many respondents reported that they are actually using altmetrics more often than in the past as they are using them the same amount.

As one anonymous community respondent explained, "There is more interest, and more availability of altmetrics. But there is also more skepticism about their value." Another customer explained their own difference in use of altmetrics as being prompted by changes to how impact is defined by the Australian Research Council and National Health and Medical Research Council.

Prediction: Personal use of altmetrics will increase in the next 2 years

Two-thirds of community members believe that their use of altmetrics will expand in the near future. Only about half of Altmetric customers believed that they too would increase their use of altmetrics. Between the two groups of respondents, around 4% thought their use will decrease in the years to come.

One community respondent said, "As altmetrics are becoming more well-known, researchers and management are wanting to incorporate them into our traditional metric reporting."

Prediction: Altmetrics will continue to grow and may see a reinvention

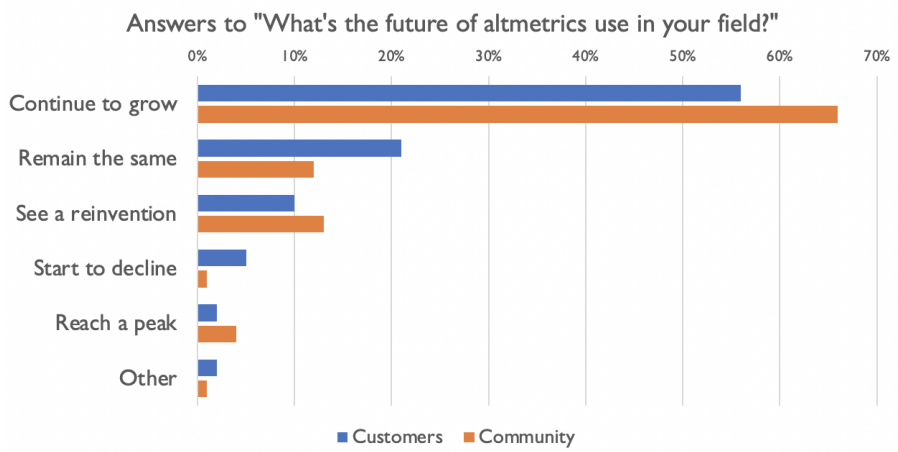


Figure 7.2: The future of altmetrics use: "Growth" was the most popular prediction.

Among both customers and the broader community, respondents overwhelmingly believe that use of altmetrics in their field will continue to grow: "Altmetrics provide data to measure open science. Since the trend towards open science continues and accelerates, altmetrics will gain more importance automatically."

Around 1 in 10 respondents think that altmetrics as a field will soon see a reinvention: "I think there may be some ways to create altmetrics, and visualize them and use them, that are not yet tapped."

Prediction: Altmetrics will be more transparent and predictive

Community members and customers alike had a lot of great ideas for future altmetrics data and products, including:

1. **Enabling better impact tracking for "long-tail" research**, sidestepping the Matthew effect[1] that influences not only citation-based metrics but altmetrics.
2. **Integration of altmetrics data into data science tools like Google Big Query**, which will allow users to do more, faster.
3. **Increased transparency and democracy** in how altmetrics are collected, reported, and aggregated.
4. **Allowing prediction of other indicators**, especially for future citations.

An overarching theme to the feedback centered around **context**. Many users simply want future altmetrics tools to be more informative and contextualized, making it easier to use altmetrics data and share it with others.

Where we go from here

In surveying the community, we have been offered a forward-thinking and optimistic vision for the future of altmetrics.

We can expect altmetrics use to continue to grow, especially among academics. The community should look to find ways to "reinvent" the field over time, perhaps by developing new product features, collecting new data in a more transparent and democratic manner, or by helping existing altmetrics users find new ways to apply the data in their daily lives (e.g. decision-making).

At Altmetric, we are eager to see how the future of the field unfolds, and looking forward to building that future alongside the community.

Liesa Ross is Altmetric's Director of Market Development for the Americas. She's enjoyed 20+ years of working with science-based and non-profit organizations to reach and serve their stakeholders, specifically in promoting the societal contributions of scientific research.

Stacy Konkiel is Director of Research Relations at Altmetric. She has written and presented widely about altmetrics. Previously, Stacy worked with teams at Impactstory, Indiana University & PLOS. You can learn more about Stacy at stacykonkiel.org.

[1] Merton, R. K. (1968). The Matthew effect in science: The reward and communication systems of science are considered. Science, 159(3810), 56-63. https://www.doi.org/10.1126/science.159.3810.56

Altmetric's mission is to help others understand the influence of research online. We collate what people are saying about published research in sources such as the mainstream media, policy documents, social networks, blogs, and other scholarly and non-scholarly forums to provide a more robust picture of the influence and reach of scholarly work. Altmetric works with some of the biggest publishers, funders, businesses and institutions around the world to deliver this data in an accessible and reliable format. Find out more at www.altmetric.com.

Contact
Altmetric, 6 Briset Street, London EC1M 5NR, UK. info@altmetric.com

Copyright
Copyright © 2020 the contributors. This work is licensed under the Creative Commons Attribution 4.0 International License CC-BY.

Cite as
Konkiel, S. (Ed.). (2020). *The State of Altmetrics: A tenth anniversary celebration.* London, UK: Altmetric. doi: https://www.doi.org/10.6084/m9.figshare.13010000

This report was created in Overleaf