

Now I am become DOI, destroyer of gatekeeping worlds

Tal Yarkoni¹

¹Affiliation not available

April 17, 2023



Now I am become DOI, destroyer of gatekeeping worlds

TAL YARKONI

READ REVIEWS

WRITE A REVIEW

CORRESPONDENCE:

DATE RECEIVED:

June 10, 2015

DOI:

10.15200/winn.142557.78481

ARCHIVED:

March 05, 2015

CITATION:

Tal Yarkoni, Now I am become DOI, destroyer of gatekeeping worlds, *The Winnower* 2:e142557.78481, 2015, DOI: [10.15200/winn.142557.78481](https://doi.org/10.15200/winn.142557.78481)

© Yarkoni This article is distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and redistribution in any medium, provided that the original author and source are credited.



Digital object identifiers (DOIs) are much sought-after commodities in the world of academic publishing. If you've never seen one, a DOI is a unique string associated with a particular digital object (most commonly a publication of some kind) that lets the internet know where to find the stuff you've written. For example, say you want to know where you can get a hold of an article titled, oh, say, *Designing next-generation platforms for evaluating scientific output: what scientists can learn from the social web*. In the real world, you'd probably go to Google, type that title in, and within three or four clicks, you'd arrive at [the document you're looking for](#). As it turns out, the world of formal resource location is fairly similar to the real world, except that instead of using Google, you go to a website called dx.DOI.org, and then you plug in the string '10.3389/fncom.2012.00072', which is the DOI associated with the aforementioned article. And then, poof, you're [automagically linked](#) directly to the original document, upon which you can gaze in great awe for as long as you feel comfortable.

Historically, DOIs have almost exclusively been issued by official-type publishers: Elsevier, Wiley, PLoS and such. Consequently, DOIs have had a reputation as a minor badge of distinction—probably because you'd traditionally only get one if your work was perceived to be important enough for publication in a journal that was (at least nominally) peer-reviewed. And perhaps because of this tendency to view the presence of a DOIs as something like an implicit seal of approval from the Great Sky Guild of Academic Publishing, many journals impose official or unofficial commandments to the effect that, when writing a paper, one shalt only citeth that which hath been DOI-ified. For example, here's a boilerplate Elsevier statement regarding references (in this case, taken from the [Neuron author guidelines](#)):

References should include only articles that are published or in press. For references to in press articles, please confirm with the cited journal that the article is in fact accepted and in press and include a DOI number and online publication date. Unpublished data, submitted manuscripts, abstracts, and personal communications should be cited within the text only.

This seems reasonable enough until you realize that citations that occur “within the text only” aren't very useful, because they're ignored by virtually all formal citation indices. You want to cite a blog post in your Neuron paper and make sure it counts? Well, you can't! Blog posts don't have DOIs! You want to cite a what? A tweet? That's just crazy talk! Tweets are 140 characters! You can't possibly cite a tweet; the citation would be longer than the tweet itself!

The injunction against citing DOI-less documents is unfortunate, because people deserve to get credit for the interesting things they say—and it turns out that they have, on rare occasion, been known to say interesting things in formats other than the traditional peer-reviewed journal article. I'm pretty sure if Mark Twain were alive today, he'd write the best tweets EVER. Well, maybe it would be a tie between Mark Twain and the [NIH Bear](#). But Mark Twain would definitely be up there. And he'd probably write

some insightful blog posts too. And then, one imagines that other people would probably want to cite this brilliant 21st-century man of letters named [@MarkTwain](#) in their work. Only they wouldn't be allowed to, you see, because 21st-century Mark Twain doesn't publish all, or even most, of his work in traditional pre-publication peer-reviewed journals. He's too impatient to rinse-and-repeat his way through the revise-and-resubmit process every time he wants to share a new idea with the world, even when those ideas are valuable. 21st-century [@MarkTwain](#) just wants his stuff out there already where people can see it.

Why does Elsevier hate 21st-century Mark Twain, you ask? I don't know. But in general, I think there are two main reasons for the disdain many people seem to feel at the thought of allowing authors to freely cite DOI-less objects in academic papers. The first reason has to do with permanence—or lack thereof. The concern here is that if we allowed everyone to cite just any old web page, blog post, or tweet in academic articles, there would be no guarantee that those objects would still be around by the time the citing work was published, let alone several years hence. Which means that readers might be faced with a bunch of dead links. And dead links are not very good at backing up scientific arguments. In principle, the DOI requirement is supposed to act like some kind of safety word that protects a citation from the ravages of time—presumably because having a DOI means the cited work is important enough for the watchful eye of Sauron Elsevier to periodically scan across it and verify that it hasn't yet fallen off of the internet's cliffside.

The second reason has to do with quality. Here, the worry is that we can't just have authors citing *any* old opinion someone else published somewhere on the web, because, well, think of the children! Terrible things would surely happen if we allowed authors to link to unverified and unreviewed works. What would stop me from, say, writing a paper criticizing the idea that human activity is contributing to climate change, and supporting my argument with "citations" to random pages I've found via creative Google searches? For that matter, what safeguard would prevent a brazen act of sockpuppetry in which I cite a bunch of pages that I myself have (anonymously) written? Loosening the injunction against formally citing non-peer-reviewed work seems tantamount to inviting every troll on the internet to a formal academic dinner.

To be fair, I think there's some merit to both of these concerns. Or at least, I think there *used* to be some merit to these concerns. Back when the internet was a wee nascent flaky thing winking in and out of existence every time a dial-up modem connection went down, it made sense to worry about permanence (I mean, just think: if we had allowed people to cite [GeoCities](#) webpages in published articles, every last one of those citations links would now be dead!) And similarly, back in the days when peer review was an elite sort of activity that could only be practiced by dignified gentlemen at the cordial behest of a right honorable journal editor, it probably made good sense to worry about quality control. But the merits of such concerns have now largely disappeared, because we now live in a world of marvelous technology, where bits of information cost virtually nothing to preserve forever, and a new post-publication platform that allows anyone to review just about any academic work in existence seems to pop up every other week (cf. [PubPeer](#), [PubMed Commons](#), [Publons](#), etc.). In the modern world, nothing ever goes out of print, and if you want to know what a whole bunch of experts think about something, you just have to ask them about it on Twitter.

Which brings me to this blog post. Or paper. Whatever you want to call it. It was first published on my blog. You can find it—or at least, you *could* find it at one point in time—at the following URL: <http://www.talyarkoni.org/blog/2015/03/04/now-i-am-become-doi-destroyer-of-gates>.

Unfortunately, there's a small problem with this URL: it contains nary a DOI in sight. Really. None of the eleventy billion possible substrings in it look anything like a DOI. You can even scramble the characters if you like; I don't care. You're still not going to find one. Which means that most journals won't allow you to officially cite this blog post in your academic writing. Or any other post, for that matter. You can't cite my post about [statistical power and magical sample sizes](#); you can't cite Joe Simmons' Data Colada post about [Mturk and effect sizes](#); you can't cite Sanjay Srivastava's discussion

of replication and falsifiability; and so on ad infinitum. Which is a shame, because it's a reasonably safe bet that there are at least one or two citation-worthy nuggets of information trapped in some of those blog posts (or millions of others), and there's no reason to believe that these nuggets must all have readily-discoverable analogs somewhere in the "formal" scientific literature. As the Elsevier author guidelines would have it, the appropriate course of action in such cases is to acknowledge the source of an idea or finding in the text of the article, but not to grant any other kind of formal credit.

Now, typically, this is where the story would end. The URL can't be formally cited in an Elsevier article; end of story. BUT! In this case, the story doesn't quite end there. A strange thing happens! A short time after it appears on my blog, this post also appears—in virtually identical form—on something called **The Winnower**, which isn't a blog at all, but rather, a respectable-looking alternative platform for scientific publication and evaluation.

Even more strangely, on The Winnower, a mysterious-looking set of characters appear alongside the text. For technical reasons, I can't tell you what the set of characters actually is (because it isn't assigned until this piece is published!). But I can tell you that it starts with "10.15200/winn". And I can also tell you what it is: It's a DOI! It's one bona fide free DOI, courtesy of The Winnower. I didn't have to pay for it, or barter any of my services for it, or sign away any little pieces of my soul to get it*. I just installed a **WordPress plugin**, pressed a few buttons, and... poof, instant DOI. So now this is, proudly, one of the world's first N (where N is some smallish number probably below 1000) blog posts to dress itself up in a nice DOI (Figure 1). Presumably because it's getting ready for a wild night out on the academic town.

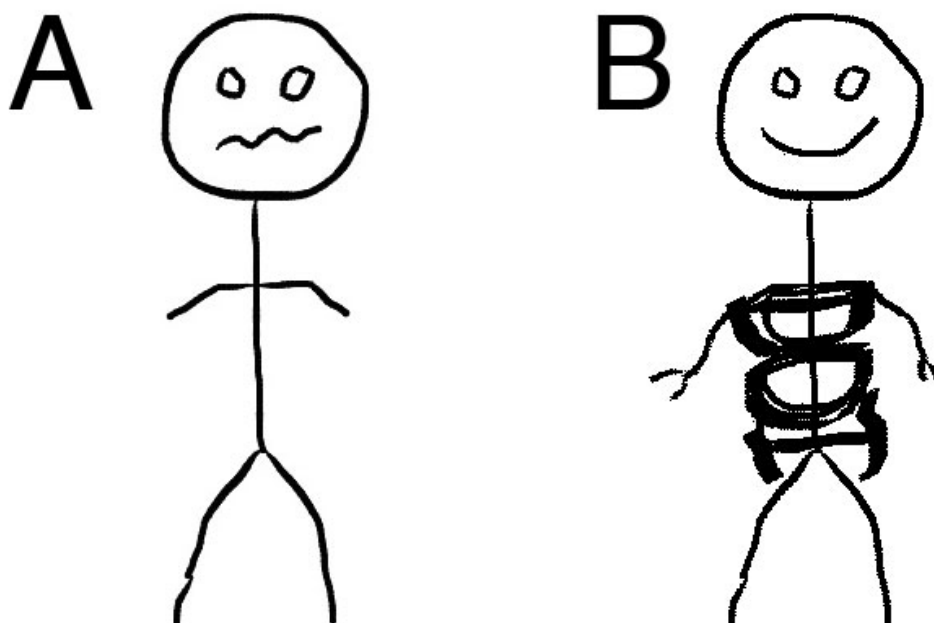


Figure 1. Effects of assigning DOIs to blog posts: an anthropomorphic depiction. **(A)** A DOI-less blog post feels exposed and inadequate; it envies its more reputable counterparts and languishes in a state of torpor and existential disarray. **(B)** Freshly clothed in a newly-minted DOI, the same blog post feels confident, charismatic, and alert. Brimming with energy, it eagerly awaits the opportunity to move mountains and reshape scientific discourse. Also, it has longer arms.

Does the mere fact that my blog post now has a DOI actually change anything, as far as the citation rules go? I don't know. I have no idea if publishers like Elsevier will let you officially cite this piece in an article in one of their journals. I would guess not, but I strongly encourage you to try it anyway (in fact, I'm willing to let you try to cite this piece in every paper you write for the next year or so—that's the kind

of big-hearted sacrifice I'm willing to make in the name of science). But I *do* think it solves both the permanence and quality control issues that are, in theory, the whole reason for journals having a no-DOI-no-shoes-no-service policy in the first place.

How? Well, it solves the permanence problem because The Winnower is a participant in the **CLOCKSS** archive, which means that if The Winnower ever goes out of business (a prospect that, let's face it, became a little bit more likely the moment this piece appeared on their site), this piece will be immediately, freely, and automatically made available to the worldwide community in perpetuity via the associated DOI. So you don't need to trust the safety of my blog—or even The Winnower—any more. This piece is here to stay forever! Rejoice in the cheapness of digital information and librarians' obsession with archiving everything!

As for the quality argument, well, clearly, this here is not what you would call a high-quality academic work. But I still think you should be allowed to cite it wherever and whenever you want. Why? For several reasons. First, it's not exactly difficult to determine whether or not it's a high-quality academic work—even if you're not willing to exercise your own judgment. When you link to a publication on The Winnower, you aren't just linking to a paper; you're also linking to a review platform. And the reviews are very prominently associated with the paper. If you dislike this piece, you can use the comment form to indicate exactly why you dislike it (if you like it, you don't need to write a comment; instead, send an envelope stuffed with money to my home address).

Second, it's not at all clear that banning citations to non-prepublication-reviewed materials accomplishes anything useful in the way of quality control. The reliability of the peer-review process **is sufficiently low** that there is simply no way for it to consistently sort the good from the bad. The problem is compounded by the fact that rejected manuscripts are rarely discarded forever; typically, they're quickly resubmitted to another journal. The bibliometric literature shows that it's possible to publish almost anything in the peer-reviewed literature **given enough persistence**.

Third, I suspect—though I have no data to support this claim—that a worldview that treats having passed peer review and/or receiving a DOI as markers of scientific quality is actually counterproductive to scientific progress, because it promotes a lackadaisical attitude on the part of researchers. A reader who believes that a claim is significantly more likely to be true in virtue of having a DOI is a reader who is slightly less likely to take the extra time to directly evaluate the evidence for that claim. The reality, unfortunately, is that **most scientific claims are wrong**, because the world is complicated and science is hard. Pretending that there is some reasonably accurate mechanism that can sort all possible sources into reliable and unreliable buckets—even to a first order of approximation—is misleading at best and dangerous at worst. Of course, I'm not suggesting that you can't trust a paper's conclusions unless you've read every work it cites in detail (I don't believe I've ever done that for *any* paper!). I'm just saying that you can't abdicate the responsibility of evaluating the evidence to some shapeless, anonymous mass of "reviewers". If I decide not to chase down the Smith & Smith (2007) paper that Jones & Jones (2008) cite as critical support for their argument, I shouldn't be able to turn around later and say something like "hey, Smith & Smith (2007) was *peer reviewed*, so it's not my fault for not bothering to read it!"

So where does that leave us? Well, if you've read this far, and agree with most or all of the above arguments, I hope I can convince you of one more tiny claim. Namely, that this piece represents (a big part of) the future of academic publishing. Not *this* particular piece, of course; I mean the general practice of (a) assigning unique identifiers to digital objects, (b) preserving those objects for all posterity in a centralized archive, and (c) allowing researchers to cite any and all such objects in their work however they like. (We could perhaps also add (d) *working very hard to promote centralized "post-publication" peer review of all of those objects*—but that's a story for **another day**.)

These are not new ideas, mind you. People have been calling for a long time for a move away from a traditional gatekeeping-oriented model of pre-publication review and towards more open publication and evaluation models. These calls have intensified in recent years; for instance, in 2012, **a special**

topic in *Frontiers in Computational Neuroscience* featured 18 different papers that all independently advocated for [very similar post-publication review models](#). Even the actual attachment of DOIs to blog posts isn't new; as a case in point, consider that [C. Titus Brown](#)—in typical pioneering form—was already experimenting with ways to automatically DOIfy his blog posts via [FigShare](#) way back in the same dark ages of 2012. What *is* new, though, is the emergence and widespread adoption of platforms like The Winnower, [FigShare](#), or [Research Gate](#) that make it increasingly easy to assign a DOI to academically-relevant works other than traditional journal articles. Thanks to such services, you can now quickly and effortlessly attach a DOI to your [open-source software packages](#), technical manuals and white papers, conference posters, or virtually any other kind of digital document.

Once such efforts really start to pick up steam—perhaps even in the next two or three years—I think there's a good chance we'll fall into a positive feedback loop, because it will become increasingly clear that for many kinds of scientific findings or observations, there's simply nothing to be gained by going through the cumbersome, time-consuming conventional peer review process. To the contrary, there will be all kinds of incentives for researchers to publish their work as soon as they feel it's ready to share. I mean, look, I can write blog posts a lot faster than I can write traditional academic papers. Which means that if I write, say, one DOI-adorned blog post a month, my Google Scholar profile is going to look a lot bulkier a year from now, at essentially no extra effort or cost (since I'm going to write those blog posts anyway!). In fact, since services like The Winnower and FigShare can assign DOIs to documents retroactively, you might not even have to wait that long. Check back this time next week, and I might have a dozen new indexed publications! And if some of these get cited—whether in “real” journals or on other indexed blog posts—they'll then be contributing to my citation count and h-index too (at least on Google Scholar). What are *you* going to do to keep up?

Now, this may all seem a bit off-putting if you're used to thinking of scientific publication as a relatively formal, laborious process, where two or three experts have to sign off on what you've written before it gets to count for anything. If you've grown comfortable with the idea that there are “real” scientific contributions on the one hand, and a blooming, buzzing confusion of second-rate opinions on the other, you might find the move to suddenly make *everything* part of the formal record somewhat disorienting. It might even feel like some people (like, say, me) are actively trying to game the very system that separates science from tabloid news. But I think that's the wrong perspective. I don't think anybody—certainly not me—is looking to get rid of peer review. What many people *are* actively working towards are alternative models of peer review that will almost certainly work *better*.

The right perspective, I would argue, is to embrace the benefits of technology and seek out new evaluation models that emphasize open, collaborative review by the community as a whole instead of closed pro forma review by two or three semi-randomly selected experts. We now live in an era where new scientific results can be instantly shared at essentially no cost, and where sophisticated collaborative filtering algorithms and carefully constructed reputation systems can potentially support truly community-driven, quantitatively-grounded open peer review on a massive scale. In such an environment, there are few legitimate excuses for sticking with archaic publication and evaluation models—only the familiar, comforting pull of the status quo. Viewed in this light, using technology to get around the limitations of old gatekeeper-based models of scientific publication isn't gaming the system; it's actively *changing* the system—in ways that will ultimately benefit us all. And in that context, the humble self-assigned DOI may ultimately become—to liberally paraphrase [Robert Oppenheimer](#) and the *Bhagavad Gita*—one of the destroyers of the old gatekeeping world.