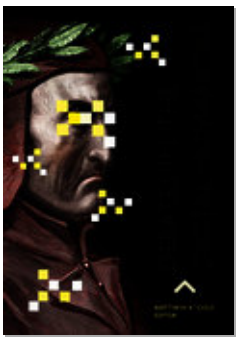
Debates in the Digital Humanities

Matthew K. Gold



Published by University of Minnesota Press

Beyond Metrics: Community Authorization and Open Peer Review

KATHLEEN FITZPATRICK

I originally began writing about peer review—its history, its present status, and its digital future—a couple of years ago, as it became increasingly clear that addressing the question of digital scholarship required a prior reckoning with the issue. I hadn't ever really intended to give it that much of my attention; but, as my colleagues and I worked on the development of the digital scholarly network MediaCommons, it kept crowding in, as it has for many digital humanities projects: at every meeting, conference presentation, panel discussion, or other venue where we discussed the kinds of work we were interested in publishing, one of the first questions we were asked was what we intended to do about peer review.

Coming to the development of MediaCommons at one and the same time from the position of being a scholar and that of being a blogger, I was of two distinct minds in responding to that question. The blogger in me wanted to wave the question off; enormous amounts of influential intellectual work are being done on the Internet today without the benefit of prepublication peer review. The scholar in me got it, however: influential as that intellectual work online may be, what is it worth within the academy's systems of accounting? What kind of value can it have, unless behind it stands some form of peer review, the gold standard of scholarly merit?

In thinking through that disjuncture in my responses, I came to understand there to exist a profound mismatch between conventional peer review as it has been practiced since the mid-twentieth century and the structures of intellectual engagement on the Internet.[1] This should not be terribly surprising; all our assumptions about scholarly communication since the eighteenth century have been predicated on the print technologies alongside and in conjunction with which they arose. These technologies operate under the conditions of certain kinds of material scarcity that, far from being hindrances, in fact give them their authority. That same relationship between scarcity and authority was operative in manuscript culture; a text's authenticity derived from the trustworthiness of its source, the hand that produced it. In

print, that sense of authority transferred from the hand to the publisher's imprimatur, which similarly served as a warranty of sorts, authorizing the book's content as being complete, correct, and stable.[2] Within the scholarly universe, however, the value of a publisher's imprimatur is not simply about that authorizing gesture but is instead directly tied to selectivity, to using the scarce resources of print in the wisest possible way. The value of a journal's name or a press's imprint varies directly with the perception that the texts it publishes are the "best"; perhaps perversely, that sense of bestness derives in significant part from the number of texts that the journal or press rejects.

The disjuncture between this model and Internet-based publishing is immediately obvious: while access to a publisher's imprimatur is intentionally limited in print, anyone with the right hardware, software, and network connection could simply publish anything online. This promiscuity is precisely what makes the open web suspect for many academics, and the usual response of those who seek to defend digital publication is to point to those online venues that employ traditional peer review and are thus equally selective as many print publishers. The problem, however, with this kind of online reproduction of the systems that govern print publishing is that the Internet operates within a completely different set of technological affordances, which have in turn helped to foster a very different understanding of the creation of authority. In contrast with print, the scarcities involved in Internet-based publishing are not material; they are, rather, shortages of time and attention—shortages, in other words, that are experienced by those who are looking for information rather than by those who provide it.

As a result, over the course of the last twenty years, the publisher-derived imprimatur declaring selectivity has gradually become less important online than the imprimatur that is conferred by community. Within the current structures of Internet-based communication, in other words, it is far less important to a given reader that some authorizing producer has declared a text to be of value—and even less important that this entity has done so by selecting this text from numerous other rejected possibilities—than it is that someone whom the reader knows, even if only by reputation, has read and recommended the text. This community imprimatur manifests itself most substantively as links that direct one's own circle of readers to valuable information elsewhere, but the relatively recent rise of the "like" button indicates the desire for and pervasiveness of lightweight community-based recommendations on the web. Liking and linking both enact a new kind of selectivity in creating a mode of community-based authorization, a selectivity that is imposed at the point of consumption rather than production.

This shift in the locus of selectivity, between the publisher's role in creating a text's authority in print and the role of communities in creating that authority online, poses key challenges for scholarship as it begins to move into networked spaces. These challenges are most pressing for the forms of scholarship that must make use of these networked environments as well as those that are attempting to

understand those environments. Fields such as the digital humanities, digital media studies, and Internet research must take the lead in reconsidering their reliance on conventionally determined and awarded imprimatur and instead consider the ways that field-based community models of authorization might more accurately reflect the innovations taking place in the field. After all, the principle of exclusion on which conventional peer review relies for its authority is part of a cluster of well-established disciplinary practices, practices that (in a most Foucauldian manner) not only create a set of accepted procedures and methods through which scholars in a field do their work but also draw the borders of acceptability—of field, of method, of knowledge—as an exercise of power. The self-policing nature of peer review, coupled with its reliance on the opinions of a very small number of usually well-established scholars, runs the risk of producing an ingrained conservatism, a risk-averse attitude toward innovation, and a resistance to new or controversial approaches.[3]

Crowdsourcing presents the potential to correct for the kind of conservatism that exclusivity can foster. Just as many eyes make all bugs shallow, as the saying has it, so many readers might make the value of a piece of scholarship more apparent. Twenty readers, with twenty different perspectives, might be able to see what two to three well-established readers cannot. Even more importantly, those twenty readers, if required not simply to respond to the work under review but to engage with one another in the process, will push one another to deeper understandings not just of the objects of their discussion but of the broader field with which their scholarship engages. In other words, crowdsourced review can improve on traditional review practices not just by adding more readers but by placing those readers into conversation with one another and with the author, deepening the relationship between the text and its audience.

It is important, on the one hand, to distinguish between the popular conception of crowdsourcing, with its sense that participation is open to "just anybody," and the kinds of crowdsourcing that can be accomplished within a preexisting community of practice, which Katherine Rowe has referred to as "our-crowd sourcing."[4] Getting the members of a scholarly field into open dialogue with one another around the evaluation of scholarship offers the possibility, as Rowe and I have explored elsewhere, of airing methodological or theoretical assumptions and biases rather than allowing them to remain covert points of contention within fields.[5] On the other hand, new forms of networked communication have raised questions about the relationship between expertise and credentials, suggesting that in many cases readers from outside our narrowly defined fields, and even outside the academy, may bring much-needed perspectives to our work. It's easy to see how open reviewing can benefit interdisciplinary work by provoking discussion among scholars from multiple disciplinary perspectives in the evaluation of a project. Beyond such obvious conversations across scholarly communities, however, many of us would benefit from discussions with readers located outside the academy. Most of our fields, and the work we produce within them, would benefit from the kinds of aeration

that exposure to other perspectives can provide; having our work discussed by readers with different forms of expertise than our own can only help us produce more robust projects.

New work being done in and on the digital not only *can* but *should* transgress the existing borders of knowledge, as these fields wrestle with new methods, new formats, and new affordances for scholarly work. The guiding principle for these fields, like that of the network technologies that facilitate them, should be openness. The Internet operates as an end-to-end network, in which any packet of data is treated exactly like any other, leaving it to the endpoints of the network to differentiate among data types, figuring out how each should be produced and used.[6] Digital fields should explore what such an end-to-end model might hold for them; if the communication network of the digital humanities were patterned after such an open principle and all scholarship were openly transmitted through it, might we find a way to allow selectivity to be exercised at the network's endpoint, with the consumers of scholarship?

This is of course the "publish, then filter" model that Clay Shirky has argued is most conducive to communication in a peer-based network, a mode of open communication that not only relocates the point of selectivity from the publisher to the reader but that also suggests, as Chris Anderson has pointed out, that network-based communication presents scholars with a radical new understanding of the notion of the "peer."[7] This understanding shifted once at the birth of scholarly communication, when the delegation of the royal imprimatur from the crown to early scholarly associations like the Royal Society of London began to press the dominant conception of the "peer" from "peer of the realm" to "scholarly peer," a colleague whose status derived less from social privilege than from earned credentials.[8] Today, that notion is shifting again, as the meritocratic notion of the peer gives way in the age of the Internet to an understanding of a peer as any node connected to a network. While I hope to avoid being overly utopian in assessing the democratizing potential of this shift, I do want us to consider how this new notion of the peer might affect scholarly communication. What would happen if we were to recognize that all members of a scholarly community—all of the nodes on our particular disciplinary or interdisciplinary network—have potentially valuable contributions to make to the development and evaluation of our work? What if we opened that network even further, to all interested readers inside and outside our fields, inside and outside the academy? What kinds of community-based filtering would such recognition make possible?

More importantly, perhaps, we also need to ask how we can get such communal processes of evaluation to *count as* evaluation in the academic sense. As academic processes of credentialing have become increasingly tied to the imprimatur granted by publisher selectivity, it has become, in turn, increasingly difficult to imagine suitable alternatives. And while it suits my inner blogger to dismiss this kind of concern—throw open the floodgates, I'd like to say; let a thousand flowers bloom;

force the traditionalists to sort the problem out for themselves—I know full well the risks that such an attitude presents for the future of the field. Scholars working in digital fields are already subject to personnel-based processes (hiring, tenure, promotion) overseen and conducted by faculty and administrators who often do not see the value of the new forms of scholarship that they are producing. The imprimatur granted by traditional peer review has often been an important recourse for those scholars whose work is controversial or defies traditional form. And yet, if that mode of peer review doesn't provide digital scholarship with the most suitable form of evaluation, the digital humanities must take control in developing a system of evaluation that does suit our needs.

Moreover, I would argue that we must develop means of articulating, clearly and compellingly, the value of open peer review—both the importance of these processes for the free and open exchange among peers and the kinds of data that such processes can provide as a means of evaluating the reception of the scholarship under consideration. As I have recently suggested elsewhere, tenure and promotion committees must find ways to understand the forms of assessment that take place around and in digital scholarship, but those of us engaging in that digital scholarship must help make our modes of assessment comprehensible to those committees, to articulate to them how the forms of review that take place at the network's endpoints function and how their results might be interpreted.[9]

In imagining ways that scholars in the digital humanities might support the robust evaluation of their work in open formats, we will need to consider a number of somewhat different cases: there are differences between the requirements for review of traditional textual scholarship published in digital environments, such as that in open access journals and online monograph series, as opposed to the review of work that does not take the form of textual argumentation, such as archival projects or multimodal work, or the review, post hoc, of DIY-published online material such as scholarly blogs. The latter two cases obviously cry out for new modes of open review, as they concern work that is nearly always published, in the sense of having been made public, without being subjected to prepublication peer review and very often without the benefit of a publisher's imprimatur. But in all three of these cases, institutional acceptance of the publications for credentialing purposes will require making visible what John Guillory has referred to as the immanent scene of judgment—laying bare for future evaluators the mechanisms by which this material has been judged by experts within the field—and finding ways to translate that immanent judgment into information that can be comprehended at the sites of review that move further and further outside the field (the department, the college, the university).

Open review online provides a number of mechanisms through which this immanent scene of review can be examined; we already have the capacity to gather a lot of data about how digital publications are received and responded to, including things like the number of page views, of visitors, and of inbound links. We have

not yet, however, found robust ways to put this kind of data to work in assessing the impact or importance of an online publication or to connect these use-metrics to peer review in a nonreductive fashion. Take something as seemingly simple as web traffic. There are many fine-grained ways of slicing this data (including entrances, exits, average time on page, and bounce rate, not to mention a metric as basic and yet amorphous as the "unique visit"), but perhaps unsurprisingly most of the readily available analytical tools are focused on the ultimate goal of "conversion": the transformation of a visit into a sale. Digital scholarly publishing will require rethinking the ways that such traffic is measured and assessed, moving from a focus on conversion to a focus on engagement—and engagement can be quite difficult to measure. A visitor to a substantive scholarly blog post, for instance, who simply snagged the post in Instapaper and moved on but then later sat and read the article at leisure would not appear to have had a significant engagement with the text when in fact that engagement could well have been far more in depth than the average. We need far better markers for that kind of engagement in order for basic traffic to become legible.

This kind of data must of course be employed in review processes with great care; many of our colleagues in the natural sciences would no doubt warn us about the dangers posed by a failure to recognize or properly contextualize the somewhat arbitrary nature of metrics such as journal impact factor. Certainly the value of scholarship in the humanities cannot be reduced to a numerical representation. We will also need to develop safeguards against certain kinds of gaming the system, such that the metrics we obtain are as nonmanipulable as possible. Nonetheless, there is a wealth of data that we're not currently putting to good use in these review processes, relying instead on the even more reductive metric of publisher imprimatur that results from traditional review.

Moreover, in addition to the quantitative data that online publications generate, there is also rich qualitative data about how they're received. The discussions on texts published in open review processes are available for examination; in networks such as MediaCommons that encourage or require commenters to sign their names to their responses, the authority that the commenters bring to the discussion can be assessed (avoiding the sneaking sense that participants in open review processes online could be "just anybody"). Similarly, inbound links, rather than simply providing numerical evidence of citation, often result from substantive discussion of the project in question. Of course, examining those discussions requires actual reading, a prospect that, as I have noted elsewhere, may be an uphill climb.[10] However, at least at the stage in which reviews for tenure and promotion move outside the subfield to the department as a whole, reviewers *should* be reading this material and, if not directly assessing the work, certainly assessing the assessment, seeing how readers directly engaged with the work have responded to it. Just as external reviewers' letters bring expert opinion to bear in the tenure process, so the results

of open peer review processes can demonstrate expert judgment as it is brought to bear on the work under consideration.

There are undoubtedly other forms of data that we should be gathering about work published online; Michael Jensen, in "Authority 3.0," presents a lengthy list of the kinds of data that might in the future be compiled in computational assessments of scholarly reputation.[11] And there are richer things that should be done with that data—the kinds of analysis for which scholars in digital humanities, digital media studies, and Internet research are ideally trained. We could, for instance, use the skill set we bring to our research to perform sophisticated network analysis on the communities that surround our digitally published texts, or semantic analysis of the discussions that they inspire. And we could, most importantly, find ways to communicate the results of that analysis in forms that would be comprehensible to members of review committees and administrators from outside the field. This solution is part computational, as Jensen argues, but also crucially qualitative and interpretive, not simply a boiling down of authority to metrics but a means of explicating what it is we value in scholarship and how we know when it has succeeded.

Scholars in the digital humanities, digital media studies, and Internet research are poised to make a profound contribution to the development of a better model of peer review, one that works with, rather than against, the Internet's open architecture and social modes of producing authority—but only if we are willing to resist the modes of discipline that we have inherited from other fields, to insist that our methods of review must more appropriately resemble and make use of our scholarly methods, and to develop the tools and modes of analysis necessary to communicate the authority generated by such review to those outside the field.

NOTES

1. For a more fully elaborated version of this argument, see chapter 1, "Peer Review," in Kathleen Fitzpatrick, *Planned Obsolescence*.

2. See Simone, especially 241.

3. See, for instance, Thurner and Hanel, whose study finds that a small number of poor referees in a conventional peer review system results in a significant reduction of the quality of work selected for publication and a tilt toward the "average." See also Dacey.

4. See Cohen.

5. See Fitzpatrick and Rowe, 139.

6. See Lessig, 44.

7. See Shirky, 98; Anderson.

8. See Biagioli, 18.

9. See Fitzpatrick, "Peer Review, Judgment, and Reading."

10. See Fitzpatrick, "Peer Review, Judgment, and Reading."

11. See Jensen, 304–5.

BIBLIOGRAPHY

Anderson, Chris. "Wisdom of the Crowds." *Nature* 2006. http://www.nature.com/nature/peerreview/debate/nature04992.html.

Biagioli, Mario. "From Book Censorship to Academic Peer Review." *Emergences* 12, no. 1 (2002): 11–45.

Cohen, Patricia. "Scholars Test Web Alternative to Peer Review." *New York Times,* August 24, 2010. http://www.nytimes.com/2010/08/24/arts/24peer.html.

Dacey, James. "Peer Review Highly Sensitive to Poor Refereeing, Claim Researchers." *PhysicsWorld.com.* September 9, 2010. http://physicsworld.com/cws/article/news/43691.

Fitzpatrick, Kathleen. "Peer Review, Judgment, and Reading." *Profession* 2011. Forthcoming.

———. *Planned Obsolescence: Publishing, Technology, and the Future of the Academy*. MediaCommons, 2009. http://mediacommons.futureofthebook.org/mcpress/plannedobsolescence/.

Fitzpatrick, Kathleen, and Katherine Rowe. "Keywords for Open Review." *LOGOS: The Journal of the World Book Community* 21, no. 3–4 (2010): 133–41.

Guillory, John. "Evaluating Scholarship in the Humanities: Principles and Procedures." *ADE Bulletin* 137 (Spring 2005): 18–33.

Jensen, Michael. "Authority 3.0: Friend or Foe to Scholars?" *Journal of Scholarly Publishing* 39, no. 1 (2007): 297–307.

Lessig, Lawrence. *Code: Version 2.0.* New York: Basic Books, 2006.

Shirky, Clay. *Here Comes Everybody: The Power of Organizing Without Organizations.* New York: Penguin, 2008.

Simone, Raffaelle. "The Body of the Text." In *The Future of the Book,* edited by Geoffrey Nunberg, 239–51. Berkeley: University of California Press, 1996.

Thurner, Stefan, and Rudolf Hanel. "Peer-review in a World with Rational Scientists: Toward Selection of the Average." *arXiv* 1008.4324 (2010). http://arxiv.org/abs/1008.4324.