Digital Technology and the Practices of Humanities Research

EDITED BY
JENNIFER EDMOND
5. Violins in the Subway

Scarcity Correlations, Evaluative Cultures, and Disciplinary Authority in the Digital Humanities

Martin Paul Eve

In January 2007, in a busy metro station in Washington DC, a violinist began to play. Of the 1097 people who walked by this violinist, twenty-seven contributed a dollar or so and seven stopped to listen. At the end of the three-quarter-hour session playing Bach, the violinist had amassed $32.17. Only one of the thousand or so passers-by recognised the busker as the world-famous virtuoso Joshua Bell who had, a mere three nights before, played the same repertoire at Boston’s Symphony Hall with good tickets going for $100 each. The violin on which Bell performed in the subway was worth $3.5m.¹

Although $30 per hour was not a bad rate of remuneration in the economic climate of 2007, the clearer point that emerges from the Bell experiment — a stunt fronted by *The Washington Post* — is that we are not very good, as a species, at identifying quality without frames of reference. As the found objects of the readymade Modernist period taught us of art: context is everything. What, though, if the same phenomena applied to scholarship? How good are we at independently judging research work, devoid of its enframing apparatus? Can we judge the music (the research) outside of the concert hall (the journal or press)?


© Martin Paul Eve, CC BY 4.0  https://doi.org/10.11647/OBP.0192.05
In the digital humanities (DH), this crux of evaluation has been entirely evident for some time. A digital historian undertaking graduate study, for instance, reported a threat that they received: ‘you will never gain a PhD doing this work’. Those working in digital literary studies are advised to publish traditional works alongside their digital projects, effectively doubling the labour required of their analogous analogue cohort. The very promise of performed new media — i.e. research artefacts that can grow and live, and that are born in the digital space — seems to re-stoke debates around print/scholarly fixity and the evaluation of ephemeral objects. Essentially, those working in the digital humanities are told, time and time again, that their work will not count. At the same time, traditional scholars often perceive the digital humanities as a ‘hot topic’ in which it is easy to gain academic employment and tenure, leading to their fear of being crowded out. Certainly, as far back as 2011, the director of the National Endowment for the Humanities’ office for the digital humanities, Brett Bobley, joked that there was a fear that DH was a ‘secret plan to replace human scholars with robots’.

Whence this conflicting sentiment? How can we understand this double logic in which DH work is at once so powerful as to crowd out the traditional humanists, while at the same time so poorly understood as to need supplementation with traditional publication? How can it be seen as both a sure-fire path to tenure but also a ‘risky thing’, as Kathleen Fitzpatrick and Mark Sample put it, to conduct digital labour

---


3 The source of this quotation asked to remain unattributed, except to note that it was spoken by a woman of incredible fire and credulity.


in the humanities? Further, in this chapter I also ask whether there is something about the broader climate at the moment in academia that fears the collapse of traditional gatekeeping mechanisms. This is linked to longer-term trends in the digital availability of scholarship and scientific publication but also to the ways in which the abundance of the digital space allows for the publication of a wider range of artefacts. For while it is possible for scholars to publish diverse types of digital artefacts, and for these to be of a high-quality, to understand the challenges of awarding credit in the contemporary age requires an understanding of the shortage of labour time for evaluation, and the necessity of frames in the evaluation of research work.

Judging Excellence and Academic Hiring and Tenure

In order to understand the broader contexts of the academy within which DH evaluation takes place, we must ask a few further questions: just how bad are we at judging whether academic work is excellent? What about within niche sub-fields? And can we tell if work closely related to our own is any good?

As with the commuters who ignored Joshua Bell in the subway, the answers are: we are very bad at judging excellence, even within niche sub-fields closely related to our own. As I have noted elsewhere, alongside many others, researchers are extremely poor at judging quality even within their own fields. This works both in terms of false negatives and false positives. For instance, in the former category, Juan Miguel Campanario, Joshua S. Gans, and George B. Shepherd examined instances of Nobel-prize winning work being rejected by

---


top journals.\(^9\) Campanario and others also note that there are many originally rejected papers that go on to be among the most highly cited in their fields.\(^10\) This is unsurprising, since most rejected manuscripts are eventually published elsewhere.\(^11\) Even more worryingly though, is that there are also instances of false positives. In 1982, Douglas P. Peters and Stephen J. Ceci re-submitted disguised papers to psychology journals that had already accepted the same works for publication. They found that only eight percent were detected as plagiarised but that ninety percent were rejected on methodological and other grounds by journals in which the material had already appeared.\(^12\) It is unclear precisely how these studies translate into the humanities disciplines, but it would not be a radical hypothesis to suggest that there may be analogies. However, despite the evidence from the above studies, most academics are usually unwilling to admit that they are unable to determine quality. Were they to do so, the entire peer review mechanism would need to be dismantled.\(^13\) Instead, there is another factor present in the understanding of the instruments through which quality is assessed in the academy: a shortage of evaluative labour.


\(^11\) See Moore et al., “‘Excellence R Us’”.


\(^13\) Despite my criticisms here, it is certainly the case that peer review may spot errors. Experts are able to question matters of fact and interpretation. They are just not good at judging the value and/or worth of work in the present.
With up to four-hundred applicants for a single academic job, hiring panels often resort to proxy measures to evaluate quality. In other words, there is insufficient labour on search committees to read and evaluate the research work of four hundred candidates, despite the fact that hiring for permanent/tenured positions represents a potential investment of several million dollars over the life of a career. While the final shortlist of candidates may have their work read, others are often eliminated by recourse to the press/journal name in which they were published, or nebulous citation measures such as the impact factor (IF). This is clearly poor academic practice that does not allow for the discrepancy between the container and its contents, and which has led to declarations such as the ‘San Francisco Declaration on Research Assessment’ (DORA), disavowing such techniques. For it is certainly the case, for example, that top university presses can publish bad books but also that low-ranking journals can contain gems. Academic freedom should entail the ability to submit one’s work wherever one wishes. However, such freedom is severely constrained by this mechanism of proxy evaluation that concentrates material rewards upon specific publication brands.

This mechanism of proxy evaluation can ‘work’ for hiring panels because publication brand correlates with scarcity, as do the applicant-to-position ratios in universities. That is to say, for example, that if it is believed that having two books with top university presses is probably achievable by around one in four hundred candidates, then the proxy works perfectly for the above hypothetical hiring scenario. In this way, publication and evaluation through scarcity proxy measures act as a symbolic economy. The currency of this economy is research artefacts, which can be traded, through hiring, promotion, and tenure panels, into a real-world material economy (jobs, pay, benefits, healthcare, pensions) for the select few.

16 ‘San Francisco Declaration on Research Assessment’, DORA, https://sfdora.org/read/
The essence of this evaluative culture is one that uses the frame of publication to judge the quality of research, usually problematically centred around a presumed single/individual author.\textsuperscript{17} It is the same type of frame that uses the concert hall to judge the violinist, and that lacks discriminatory power when that same violinist appears in the subway. While it may be true, as Kathleen Fitzpatrick suggests, that ‘we must be willing to engage in the act of judgment ourselves’, we must also acknowledge the difficulties we face in undertaking such acts.\textsuperscript{18} Importantly, though, this culture is also one that confers value upon specific media. University presses, for instance, publish books and journals. When ‘university presses’ are, then, the way in which hiring, promotion, and tenure panels make their decisions, there is an implicit underlying constraint of the valid forms that may be framed for evaluation through such proxy measures. Further, there is the matter of the continued belief in the efficacy of peer review, despite evidence to the contrary, which is linked to a reinforcement of existing media types. For example, if the labour of peer review is itself a type of service practice on which academics are assessed, the motivation to review for a high-profile press — whose brand will once again help with career advancement — is stronger than the motivation to review for radical/new publishers. This then reinforces the types of media that those traditional press entities publish, since peer review must be attached to particular objects and media types. In other words, there is a strong circularity of incentives for both authorship and peer-reviewing practices that severely constrict change in the type of media through which academics are assessed.

Finally, it is also necessary to pay attention to disciplinarity as a constraining factor in the evaluative cultures of university hiring. Disciplinary segregation, as Samuel Weber has charted it, is a way of


amplifying authority through the delineation of a sub-field over which one may preside. In other words:

[in] order for the authority of the professional to be recognized as autonomous, the ‘field’ of his ‘competence’ had to be defined as essentially self-contained […] In general, the professional sought to isolate in order to control.

and

[the university, as it developed in the latter half of the nineteenth century, became the institutional expression and articulation of the culture of professionalism.\textsuperscript{19}]

But, as the old advice for graduate students used to run in the UK, while English programmes churned out competent critical theorists, there were no critical theory departments; one had to be a literature scholar. Even within the realm of the digital humanities, though, this urge towards bounding and containment in the name of intellectual authority is a vocal debate.\textsuperscript{20} That said, it is frequently recognised that the digital humanities is an interdisciplinary space, even when it is difficult to define this term.\textsuperscript{21} So, while citing an unpopular figure in digital humanities circles, it remains true that ‘being interdisciplinary is so very hard to do’.\textsuperscript{22} Indeed, the tendency of interdisciplinary methods to rest upon a dominant home discipline, while legitimating themselves as being ‘interdisciplinary’ through reference to an exotic other discipline, is omnipresent. But the sites of authority through evaluation in universities are primarily grouped around traditional disciplinary categories that can feel threatened by digital incursions.


The Diverse Media Ecology of Digital Humanities

There are, in the taxonomy that I have sketched above, three conjoined and self-reinforcing elements of the academic evaluative cultures around research:

- a desired scarcity correlation between the research artefact and the position;
- a frame for evaluation that denotes scarcity, that is media specific, and that saves evaluative and reviewer labour;
- a set of disciplinary norms and agreements about which frames (in point 2) best denote comparable scarcity (in point 1).

The digital humanities, or in some cases just the digital, pose threats to a number of these evaluative cultures.

To begin to unpick this, consider that digital dissemination in general is causing problems for the scarcity correlation. In previous eras, the scarcity correlation was obtained through material print scarcity. That is, before digital dissemination was possible, a limited page budget with comparatively high printing costs per page came together to enforce a condition of scarcity. The digital environment changes this. In the world of the digital the vast majority of costs are shunted into the costs required to produce the first copy, which is still far from negligible in the academic publishing space (labour functions and estates costs include: typesetting, copyediting, proofreading, platform maintenance, digital preservation, identifier assignment, report generation, accountancy, legal, property, and equipment), while the costs of producing subsequent copies become almost zero. By decreasing unit cost and also by moving different forms of labour onto authors, as Matthew G. Kirschenbaum has recently noted, the print scarcity that previously underpinned the scarcity correlation for quality begins to collapse.\(^{23}\) As journal articles and books — the previous media of print scarcity — become digital in their production, their scarcity function, which was always an economic function, is degraded. This is a little like the dropping of the gold standard as a way to measure the value of currency. Except, in the

---

case of academic hiring, the belief in the value of the artefact, decoupled from any non-imagined scarcity, does not seem sufficient to continue.\textsuperscript{24} In this way, at the heart of the digital’s possibilities of infinite near-zero-cost dissemination lies the antithesis to the scarcity that has been used as a hiring proxy until now.

The practices of those working specifically in some form of the digital humanities, though, pose a set of additional challenges not only for the scarcity correlation but also for the frames of evaluation. In common with other scientific disciplines, the rise of the need to disseminate diverse forms of quantitative and qualitative data, software/code, and interactive artefacts within DH breaks the conditions of scarcity in a very particular way. For it is not precisely that such artefacts (the ‘project as basic unit’ as Anne Burdick et al. put it) are not scarce.\textsuperscript{25} Whether it is the Digital Library of the Caribbean or the Манускрипт project of Udmurt State University and Izhevsk State Technical University that is under discussion, these projects are often unique; the ultimate form of scarcity. The same could be said, of course, of conventional academic books and articles, which are supposed to be unique in their original contributions to knowledge. Yet, books and journal articles are treated as comparable units of currency, while data, code, and interactive exhibits, in their uniqueness, are usually treated as though they were incomparable. In this way, there is a belief in the comparability of artefacts and the way they can be reviewed based on the shared, or otherwise, characteristics of their media form (digital vs print).

This supposition of (in)comparability is predicated on the belief of the uniform (or otherwise) nature of peer review that is tied to media form. Although the peer review’s gatekeeping process is usually kept hidden due to concerns about anonymity and the freedom to speak truth to power, the furore around \textit{PLOS ONE’s} reduced threshold of evaluation for publication illustrates this anxiety. In the \textit{PLOS ONE}

\textsuperscript{24} There are some challenges with the divide that I am here drawing between a digital abundance and a material scarcity. Since labour is itself scarce and tied to material economic scarcity, and since there is labour in publishing, there remains a real non-imagined scarcity even in the digital. For more on this, see Martin Paul Eve, ‘Scarcity and Abundance’, in \textit{The Bloomsbury Handbook of Electronic Literature}, ed. by Joseph Tabbi (London: Bloomsbury, 2017), pp. 385–98, https://doi.org/10.5040/9781474230285.ch-022

model, work is appraised on its technical soundness rather than on its novelty, originality, or significance. This mode of peer review is designed to encourage replication studies and the publication of negative results, aspects that are also of interest to many data-driven sub-fields in DH, such as stylometry and sentiment analysis. However, as noble as its scientific purposes may be, PLOS ONE’s altered review model causes substantial problems for hiring and evaluative proxies. How should the name PLOS ONE be viewed alongside Nature or Science? In deliberately lowering its scarcity threshold in the name of good science, PLOS ONE asked the academic community to examine its own processes for evaluation. In making itself unique as a mega-journal with this threshold, the brand of the journal has been altered. Yes, PLOS ONE was, at the time, itself scarce as the only entity of its type, but the quality threshold was not determined as comparable with other outlets and so the scarcity function was eroded. The unit of currency became non-exchangeable.

A similar problem occurs in one-off DH projects. Uniquely scarce, of course; these artefacts contribute to the diverse media ecology of the digital humanities. Yet, their very uniqueness, while being scarce, is non comparable. That is, because they are one-offs, developing standards for comparability is a disproportionate activity in terms of labour time, that does not fulfil the second characteristic I outlined above. In other words, to evaluate the artefact, as itself, rather than through a proxy of presumed-uniform review, has no labour-saving function. This is why such unframed projects begin to cause anxiety among those who have come to rely on the proxies that they believe denote comparable scarcity.

This is, in part, why we have seen the emergence of documents such as the MLA’s ‘Guidelines for Evaluating Work in Digital Humanities and Digital Media’ or the AHA’s ‘Guidelines for the Professional Evaluation of Digital Scholarship by Historians’. However, these guiding

documents often struggle to fulfil the ‘needs’ of hiring committees. That is, in asking for respect for medium specificity, alongside the requirement for the engagement of qualified reviewers — or, as Sheila Cavanagh puts it, by asking for consideration of the ‘complicated factors in the world of digital scholarship needing attention’ — such guidelines do not alleviate the labour shortage of the search panels nor do they provide a uniform comparability mechanism for scarcity.\textsuperscript{27} Although these go unacknowledged, since most panel members do not wish to admit that they need recourse to such proxies, the continued fetishisation of print (for its scarcity) and the desire for hidden, yet claimed, uniform and comparable media-constraining gatekeeping practices, all highlight why it remains difficult for the proliferation of new digital artefacts to be easily integrated within conventional hiring mechanisms.\textsuperscript{28}

Strategies for Changing Cultures: Disciplinary Segregation, Print Simulation, and Direct Economics

The diverse media ecology of DH poses a threat to the first two areas in which hiring panels and accreditation mechanisms operate: in the abundance of its artefacts, the digital disrupts scarcity, while in the uniqueness of its outputs, it defies the comparability of proxy frames. The final area in which DH causes anxiety is in its inter-/multi-/trans-disciplinary nature. The challenge that DH creates in this final space is one of both evaluation and authority. In the first case, conventional hiring panels often struggle to evaluate part of a DH project; that is, the digital part. A lack of statistical knowledge among members of a search committee can also cause trouble here for certain types of DH practice. The authority challenge that is posed here is an unseating from their thrones of those with insufficient digital knowledge to carry out evaluation. This is the same challenge that other fields, such as religious history, can face: to be hated by both theologians and historians. To have created a ‘discipline’ usually means that one understands the


\textsuperscript{28} I am aware that there are other good reasons to stick to print for long-form reading. However, in the assessment domain, it is the scarcity that is valued.
evaluative requirement within that space. The practices of DH, which can intrude upon any of the conventional humanities disciplines, are challenging to those at the top of the pyramid since they suddenly find that they are not masters of their own kingdom. The work purports to be in a subject area that is recognisable to them but they know neither how to evaluate it nor how to test the research for relative soundness. When a discipline cannot evaluate work that purports to be within its own subject area, it faces a crisis. Hence why DH poses such a threat. Max Planck once famously put it that science advances ‘one funeral a time’. Since disciplines are self-reinforcing spaces, though, it is not even clear that this is the case; value systems are absorbed and internalised by those who travel through the academic ranks.  

One of the strategies for avoiding this interdisciplinary threat has been to establish and strengthen specific DH departments. At the time of writing the most recent example of this was at King’s College London where the department of Digital Humanities advertised for eight permanent, full-time posts (tenured equivalents) ranging from lecturers up to full professors. Likewise, the School of Advanced Study at the University of London is seeking a candidate to lead a new national centre for digital practice. By demarcating the space of expertise to a specifically digital domain it is possible to pursue digital practices and to hire staff members in ways that do not appear to compromise disciplinary expertise or authority. On the other hand, this also leads to a potentially problematic ‘siloization’ of digital expertise and the merely static reproduction of other disciplinary norms on which it is often the purpose of DH to intrude. For example: what use are authorship attribution technologies if nobody who defines themselves as a traditional literary scholar pays any attention? What is the point of spatio-temporal mapping approaches and GIS techniques if they cannot be used to inform other disciplinary cultures? From a research point of view, the banishment of DH to its own departmental area is a problematic move.

That said, DH as a departmental space makes sense from the economic perspective of teaching. Such programmes, which can often

---

promise transferable practical skills training and general computational
thinking, recruit well; although they have also come under fire for

The general difficulty, though, is in the intellectual breadth covered by the single seemingly simple word: humanities. Some humanists have more in common with mathematics than with literature, while others are more akin to social sciences. In the disciplinary segregation of DH then, at the same time as providing for a broader perspective and harnessing the benefits of a wide set of views that transcend any single discipline, problems of an incoherent intellectual space can emerge. It is unclear, though, at least to me, whether computational approaches are enough to bind together such otherwise disparate fields of practice in perpetuity. For the sake of binding together these fields into a space of intellectual authority, we may see a set of changes — positive or negative — around disciplinary coherence.

A similar separation of DH is evident in the proliferation of new publishing venues for the field.\footnote{Without veering too far into the ‘defining DH’ genre, see Alan Liu, ‘Is Digital Humanities a Field? — An Answer from the Point of View of Language’, \textit{Alan Liu} (6 March 2013), http://liu.english.ucsb.edu/is-digital-humanities-a-field-an-answer-from-the-point-of-view-of-language/ for more on the use of the term ‘field’.} That many of these are still journals (the \textit{Journal of Digital Humanities}, \textit{Digital Humanities Quarterly}, and \textit{Digital Scholarship in the Humanities}, are just three examples) speaks to a deep understanding among many digital humanists of the challenges of evaluative framing and media outlined above. Even though hiring panels could delegate evaluative authority to a DH community that somehow gatekept projects, by adhering to the understood media-form of the journal article, research outputs become an exchangeable currency in diverse disciplinary settings. Similarly, book chapters are a recognisable form that play into the long history of the codex, but that are, in digital form, mostly a simulation of print. Such a simulation is effective since it appears to be a simulation of the form of material scarcity that was previously inherent within print. In other words, even while the greatest costs continue to inhere in selectivity, print simulation
is maintained so that the illusion of scarcity economics can be preserved within our systems of evaluation.

This notion of the simulation of other forms that DH has had to adopt is profitable. Indeed, many scientific disciplines also feel this pain of separation between the research outputs they produce and the work they conduct. This is why the recent practice of data sharing has simultaneously become both a welcome activity and a contentious one. Billing the sharing of data as better for replication and verification is an easy argument to make. Without it, journal articles are just descriptions of work without the underlying work itself: a print simulation of non-print activities. On the other hand, very few scientists would consider submitting a dataset to any evaluation exercise as the work itself.32 The same goes for software and toolsets in the digital humanities; as Susan Schreibman and Ann M. Hanlon found, there is a ‘relationship to [the] scholarship’ of software in which many creators feel their work to be a scholarly activity, even while claiming more distant publication benefits.33

The final frame to which DH can, and does, resort is to bypass the symbolic economy entirely and move to hard currency: cash. DH is a relatively successful field in the space of research grants. Sheila Brennan addresses this: it is possible to ‘let the grant do the talking’; that is, at once to allow the fact that DH attracts money to be itself a criterion for evaluation, but also to use the accountability and documentation practices to produce an archive of creditable narrative statements around a project.34 Given that all systems of evaluations are economies, the cry of ‘show me the money’ can ring loudly. Yet, this is not likely to endear DH to traditional humanists, and it is not clear that DH will itself be spared the axe when the time comes.

In this chapter, I have explored how and why various systems from peer review and aggregation, to ‘container-level’ evaluation,
remain extremely limited, and yet are still in use. I have also made the case that all such systems of evaluation are economic in character. In turn, I have examined how the digital humanities field poses a set of challenges to the three principles of academic evaluation that I have outlined. It seems to me that it is very difficult to change the academic contexts of evaluation; they are complex social constructs (which is not to say that they do not have definitive real-world effects), not fixable technical realities. This gives a set of rationales for why DH continues to adopt publication practices that can be brought into harmony with such demands for substitution and exchange. While Lisa Samuels and Jerome J. McGann write of deformance, publication practice — for reasons of evaluation — remains in the realm of conformance, and will continue to remain there until we build our own disciplinary spaces. These too, over time, will solidify their evaluative cultures and become unyielding to, and impenetrable by, new practices. In the meantime, listen for violins in the subway when next you ride.

Bibliography


5. Violins in the Subway


Klein, Julie Thompson, Interdisciplining Digital Humanities: Boundary Work in an Emerging Field (Ann Arbor, MI: University of Michigan Press, 2014), https://doi.org/10.3998/dh.12869322.0001.001


‘San Francisco Declaration on Research Assessment’, *DORA*, https://sfdora.org/read/


