Creating scholarly knowledge in the digital age

Ziyad Marar

- Open movements focus on the consumption of information but neglect to focus on its mode of production.
- In a world where increasing amounts of information and knowledge are available, what matters is the ability to create and attend to that which is good and relevant.
- In the world of scholarly knowledge 'good' means not popular but authoritative. We must not lose sight of the values and mechanisms that sustain authority in favour of the blunt and measurable traffic of information as commodity.
- Some form of pre-selection and quality control of claimed new knowledge is therefore required and this is what publishers of journals and books provide.
- Selection mechanisms necessarily differ from discipline to discipline because scholarly knowledge is not homogeneous and the routes to it are various.
- Knowledge in HSS is more closely linked to the individuals who have produced it than in the large team-based projects of the natural sciences.
- Early career authors need to build their reputations and thereby their claims to authority; publishers have a crucial role to play in this process.

A great theme of our digital age is around openness with a corollary emphasis on a 'free culture'. The untimely death of hacktivist Aaron Swartz has led to only the most vivid flare-up of this apparently democratic call to arms. The examples hardly need rehearsing: thanks to the rise of Wikipedia, the blogosphere and Twitter, YouTube, the open access and open data movements, and now MOOCs (massive open online courses), we cannot doubt that an extraordinary technology-driven revolution towards the frictionless transmission of ideas is under way.

As we watch the disruption of various creative, cultural and knowledge industries from music to journalism, and increasingly publishing and higher education, we are retold this motto with the warning that anyone

who just doesn't get it should get out, or wait to be swept out, of the way. The verdict of commentators at the 2013 World Economic Forum at Davos was that the whole higher education system is under dramatic change and many institutions will, and should, fail. [1]

Undoubtedly we have all benefited wonderfully through access to information and knowledge that was previously inaccessible. And there is indeed something democratising about unfettered access to vast stores of information alongside immensely increased levels of participation in our digital culture. The titles of best-sellers like Clay Shirky's *Here Comes Everybody* and James Surowiecki's *The Wisdom of Crowds* capture the spirit. Shirky in fact introduced the idea of 'publish, then filter' long before it appeared in his 2008 book. As far back as 2002, he told a BBC audience the following:

> The order of things in broadcast is 'filter, then publish'. The order in communities is 'publish, then filter'. If you go to a dinner party, you don't submit your potential comments to the hosts, so that they can tell you which ones are good enough to air before the group, but this is how broadcast works every day. Writers submit their stories in advance, to be edited or rejected before the public ever sees them. Participants in a community, by contrast, say what they have to say, and the good is sorted from the mediocre after the fact.[2]

It seems to me, however, that there is a missing, or at least under-reported, aspect of these debates. The consumption of information obviously only makes sense as a sequential step after its production. Yet an overriding focus on how knowledge can be consumed without constraint can obscure from view, and possibly distort, how it is produced in the first place. And here I am talking about novel or significant contributions to scholarly knowledge. Wikipedia by contrast is expressly designed not to introduce original ideas; it does a different job and does it extremely well. But how does original, high quality information and knowledge get produced to start with? And might some features of the new media paradigm impede that creation in some way?

The first thing to say is that as data pour in without the right filters, attention becomes the scarce resource – along the way the emphasis shifts from careful, considered, thoughtful, deep work to high impact newsy items vying for attention and popularity. Where does authority feature in this landscape? Here is LSE law professor Conor Gearty on his own writing experience:

> The old-fashioned hard work – quiet; library-based; thoughtful – that made the writer/speaker an expert in the first place gradually drifts off the daily agenda. At first because of time constraints and then – well – because it's boring, like returning to decaff coffee after an espresso. Twitter/Blog erodes our confidence in the deeper stuff without which we would never have become experts in the first place.[3]

What is happening to the incentive structures – currently secured through the reputation conferring mechanisms of significant journal and publisher brands – that enable the effort required to create authoritative knowledge claims alongside the concentrated attention needed to consume them? The speedy shifts we are seeing in the digital age, while delighting us as consumers, might well, as Gearty says, erode existing mechanisms through which people become experts in the first place, without offering an adequate alternative.

The related but more profound problem on which I would also like to focus in this essay is that the free and open movements and hacker communities tend to presuppose that knowledge is to some extent commoditised. This perspective understandably encourages sharing and discourages hoarding. If the world of data is a discovery process, a bit like mining, then we should truffle out nuggets of knowledge and circulate them as rapidly as possible to where they are needed most.

No wonder the digital debate focuses on how to make that communication process yet more frictionless and the vested interests more redundant. In the new world why would we bother with publishers, learned societies, journals, even universities at all? Let the crowd do its work. As I will argue

in the rest of this essay, this view underestimates the hugely divergent ways knowledge claims are produced in different fields.

One size does not fit all

The rush to openness and its tendency to elide key differences between types of knowledge is not just a feature of the supposed wisdom of digital crowds, and their hunger for information. It is sometimes embedded within the scholarly establishment itself. Take one obvious example: when the Research Councils UK (RCUK) originally implemented its mandate requiring open access publishing from 1st April 2013 it did so equally for all and any scholars who receive funding from research councils – on the assumption that one size fits all whether in sociology or medicine. This undifferentiated approach, for example, mandates publication under a CC-BY licence (the most permissive licence enabling derivative reuse of the author's work). It was greeted with enthusiasm by scholars in the Biological Sciences and deep concern from many in the Humanities and Social Sciences. [4] Why should this be?

In general it should go without saying that ideas, information, data and knowledge cannot be lumped together into an undifferentiated mass. While light bulbs or pound coins are, in economic terms, fungible – that is to say interchangeable without loss – in many cases this cannot be said for high quality intellectual property. The pub-owner who buys books by the yard to adorn the walls and make a cozy club-like atmosphere treats a random yard of books as equal to any other. For the rest of us there is, or ought to be, a huge difference: there are good and bad ideas, interesting and trivial ones, subtle and simple, prosaic and poetic, technical and commonplace, all produced with varying degrees of care and authority and published as outputs ranging from books, chapters, journal articles, conference papers and posters, through to newspaper and magazine journalism (long and short form), essays, blogs and tweets. Layered on to this in the realm of academia we see a diverse ecosystem of scholarship with varying norms and working practices in disciplines ranging from high energy Physics to History. The differences should be obvious but in

the 'information wants to be free' environment they start to sound like undifferentiated bits and bytes that just need to be uncaged.

New ideas in this homogenising view can tend to be seen as found rather than made, and at its most simplistic can be caricatured as the sculpture that is revealed merely by hacking away the extraneous stone. While this view is a crude distortion of reality and understates the hugely creative process that is involved in creating science, it does indicate something about the varyingly fungible nature of what is ultimately produced.

One can see how this view can be reinforced. Natural scientific knowledge claims that work to some extent become exchangeable: facts, to a degree, disconnected from their original authorship, in order to become subsumed into the work of future scientists. Like Newton standing on the shoulders of giants (and then Einstein on the shoulders of Newton if you like) the new invention is subsumed within the old, and succeeds by flowing into a settled stock of knowledge. And while the scientists concerned may become legendary names, in their own right, there is no scientific reason to read their original works. One can extract their innovations and improve on them without needing to go back to the original knowledge claim. The analogy in economic terms is of the commodity which has full or partial fungibility; that is, the market treats its instances as to some extent equivalent with little regard to who produced it. As Karl Marx put it 'from the taste of wheat it is not possible to tell who produced it, a Russian serf, a French peasant or an English capitalist.'[5]

In this way, once an original claim reaches the status of fact, it can be circulated uncontroversially to provide more shoulders for others to stand on. No one disagrees that the speed of light is around 186 thousand miles per second or that the definition of energy is mass multiplied by the speed of light squared, and, now that Einstein has handed his discovery to the world, $E = mc^2$ has lifted free of his original 1905 paper and can be used as interchangeably as a pound coin. This view of discovery which settles arguments is of course challenged in many scientific fields but has become so established as an ideal that some physicists are genuinely worried that they will have nothing to do after the discovery of the Higgs boson

completes the standard model. As Professor Marc Sher, who has devoted his entire professional life to theoretical description of the Higgs Boson put it 'Now, we're like the cat who has stalked a mouse for 35 years', he continued. 'We finally catch the mouse … and now we're wondering what to do with it.' [6]

Of course I am simplifying the processes of scientific creativity. In many scientific fields there is plenty of room for unsettled, disputatious, controversial debate, and many fields don't settle very well at all. Even so when we move from analysing the natural to analysing the human world we complicate the story at a much deeper level. This is because in the Humanities and Social Sciences facts and values blur and the data under scrutiny are particularly unruly. The philosopher Bertrand Russell once observed 'the fundamental concept in Social Science is Power, in the same sense in which Energy is the fundamental concept in physics.'[7] And while physicists can define energy with enviable unanimity the same cannot (and will never) be said for sociologists exploring the nature of power.

In the Humanities and Social Sciences, knowledge claims are highly contextualised by the unruliness of the phenomena they seek to explain. The reason that sociologists can't agree on the nature of power in the way that physicists agree on the nature of energy is not because they aren't smart enough, nor is it that they need the crowd to come in with its assessment. For one thing the outputs of social analysis can change the very nature of what is being analysed: this is known as the problem of reflexivity.[8] This is not true, in the same way, of the natural world. In addition, the very concept, while incredibly relevant and important to our lives, is messy and ill-structured. Many scholars are dealing with what have been called 'wicked problems' which don't submit in an orderly way to scientific analysis.[9]

It is relevant here to distinguish 'wicked' from 'tame' problems in the digital age because it is only in the case of the latter that the wisdom of crowds or automated mechanisms of assessment can more aptly provide good solutions: as programmers like to say 'given enough eyeballs, all bugs are shallow'. But a bug in a piece of software is a tame problem,

the nature of power in society is not. Tame problems like estimating the height of a mountain, like fixing a computer programme, will more easily respond to a crowd-sourced, 'publish, then filter' solution, as contrasted with what political scientist Robert Horn called 'social messes' which as he says are so 'resistant to analysis and, more importantly, to resolution'[10].

For this reason a lot of important theoretical work does not issue in law-like generalisation, clear-cut prediction and fungible claims. An expert analysing messy, wicked systemic phenomena like terrorism, inequality, well-being, crowd behaviour, democracy or the French Revolution is able to gain influence by having taken the trouble to apply rigorous care and a familiarity with the scholarly literature, which in turn gradually comes to mark this analysis out as worthy of attention.

Rather than standing on the shoulders of giants these thinkers are more like a network of astronomical bodies with differential gravitational force orbiting around each other. The more gravitational force – let's call that force authority – the more the idea has influence over others. So with the concept of power we have interventions from Gramsci, Foucault, Giddens and Steven Lukes to name just four. These innovators (or even innovating synthesisers) produced their knowledge claims under specific conditions. And as those claims took hold they were themselves picked up by qualified critics engaged with the same problem domain and thanks to whom their reputations were secured over time. They and their critics published papers and books, talked at conferences, in order to rework authoritative predecessors and debate credible contemporaries. But for every one of them we have heard about there were many, many who tried and failed to build their own action at a distance (Russell's own book on Power, quoted above, is not widely cited these days). Authority, prestige, reputation, credibility (pick your term) in a given domain is crucial to the production of scholarship while only being achieved by a minority.

Scholarly knowledge should not be seen as an undifferentiated lump because the values and mechanisms for innovation common among physicists differ from those of sociologists, economists, historians or

anthropologists. Each of those disciplines among others has its distinctive way of creating contributions that inform debates and its varying relationship to the concept of authorship. At one end of the spectrum scientific and technical knowledge is increasingly being produced by large international teams publishing the output of a hugely funded research project and focuses less on individual authorship (of a single authored article or a book). Many scientific articles are published these days with hundreds if not thousands of contributors.

This disconnection from authorship and consequent increased fungibility of knowledge claims in parts of the natural sciences makes them easier to trade in the attention economy, and should be contrasted to some degree with ideas in Humanities and many of the Social Sciences, which are far less separable from their authors, producing ideas (in articles or monographs) whose most valuable qualities are debased by the process of information commoditisation.

Authority, authorship and publishing

It is clear then that much scholarly knowledge, especially in the Humanities and Social Sciences, does not have a fact-like, or at least agreed upon, commonsensical quality and will not settle down in a way that can be parleyed easily into interchangeable nuggets of information without considerable loss. Nor will its importance seem obvious immediately. Rather, the significant work usually builds its influence over time and is often diffuse in its impact. More so than in the natural sciences it is created, recreated, contested, forgotten, reinvented, developed, distorted, amended by people with varying degrees of expertise and who, in winning arguments, build their credibility further. The crucial variable in this complex interaction is less to do with discovery and all the more dependent on the credibility, authority, or expertise of the author built up over time and standing as a proxy for underlying values to do with effective commitment to a discipline and a community. This fragile ecosystem depends on the many filtering and enabling mechanisms provided by publishers.

Speaking as a publisher, it is undeniable that many of the traditional functions around the dissemination of ideas have been increasingly displaced or transformed. We are all publishers now that new technologies have replaced many of those original barriers to entry. I should stress that the dissemination role of publishers is nonetheless still relevant even if we now talk of technology platforms and meta-data more than printing presses. But my purpose here is not to explore the role of publishers as disseminators. It is our role as enablers of quality knowledge production, by helping scholars secure their reputations, that I want to highlight and that has been obscured in the debate. It is easy to forget that a category of the most effective publishers in this regard are the learned societies, from the American Psychological Association to PRIO (The Peace Research Institute, Oslo), whose publishing activities will often provide the disciplines they serve with a crucial mechanism for career development and subsequent impact.

Of course once an author is well-established they can go to alternative sources to increase their impact, whether by blogging, tweeting, publishing in mainstream news media and so on. But how will they build that reputation, authority and prestige to start with? How will they *become* expert and build a robust reputation that stands the test of time? They currently do this by finding publishers or journals which select and shape their monographs or articles on some level, authorising and preserving their ideas so they can be introduced into the community. Most submissions do not make it through this filtering process (whether peer review or the slush pile) and most that do are neglected (deservedly or not), but some go on to create over time those gravitational fields of scholarly force. People need an authored version that has been selected, shaped, refined and validated in various ways so that it can have better claims to take its place in the network. This process too is contested and flawed but is better than a free-for-all.

This is not to say that *existing* publishing or journal brands need to survive. Leading publishers and journals which currently provide this kind of certification may fall by the wayside if they do not keep up with the times. But they will need to be replaced with equally reputation-

enhancing publishing brands if new scholars are to make their names. And these new brands, I predict, will need to confer authority through familiar mechanisms of pre-publication peer review for articles and equivalent academic filtering for books. Without those checks in an open environment we may have secured mechanisms for knowledge sharing and consumption, but we will have weakened the possibility of authoritative knowledge claims for the future, especially those made by early career authors.

There are moves afoot, thanks to the digital revolutions, to try and create new mechanisms for reputation building: article level metrics, usage factor measures and so on. [11] It is understandable that we should want to find alternatives to the much fetishised hallmark of scholarship known as citation indices and impact factors. The problem with these more automated approaches is that they tend towards eliding *popularity* and *authority* and that these have different underlying dynamics in the less tame and messier ends of the scholarly spectrum. Popularity measures might be sufficient for the business of making music; they are not for the business of making all scholarly knowledge. Impact factors, for all their many faults, at least address this distinction and are not distracted by measures of popularity. They try to preserve the concept that academic reputation lies in the careful eye of the *qualified* beholder.

A good reputation is one of the key rewards for committing time and effort to scholarship. It is a hard-won route to enabling a voice to be heard in the babel of other voices, and ought to be quite different from popularity and celebrity. The crowd and automation might help with measures of popularity but deal less easily with authority. Or more precisely when gauging the height of a mountain, popularity and authority may converge – the crowd may source a wise incrementally more adequate solution – but when dealing with a wicked problem the crowd may become an impediment. As the technology theorist and author Tom Chatfield helpfully puts it (personal communication), it is a category error 'where people are taking one sense of the word "authority" (arriving at a decent empirical answer to a clearly defined tame question, which online crowds are pretty good at) and then

mistakenly applying it to all other senses of the word "authority", even in fields where what it means to be authoritative is quite different.' As he goes on to point out, this category error is 'mirrored in the tendency of the attention economy to value all knowledge only in terms of its effectiveness at commanding attention, or becoming grist to the mill of aggregation'.

John Ruskin once observed that 'quality is never an accident, it is always the result of intelligent effort'. This is as true for commentary as it is for the pieces being commented on, and we need to ensure that there are incentive mechanisms which enable intelligent effort to be used to assess knowledge claims. For this reason peer review, which incentivises authoritative reviewers to read and comment with care and a sense of responsibility to the discipline before publication, still has a claim to the Churchillian defence of being the least worst system: in need of improvement, certainly, but far from broken as many claim. Chatfield goes on (personal communication) to summarise the problematic thus:

• Considerable time and effort are required for work in many fields if one is valuably to contribute to those fields
• A vigorous community of those investing such time and effort is also required for valuable work to be sustained
• Supporting and sustaining this community in turn requires dedicated 'enabling' mechanisms
• These enabling mechanisms should centrally involve the (1) filtering (2) publication/dissemination, and (3) debate of work in these fields
• The best such enabling mechanisms are not algorithmic or automated, but performed by members of this community
• The proper enactment of these mechanisms requires time, effort and a sense of responsibility to the discipline
• This brings us back to the start: there is no substitute for human time and effort in many fields.

What alternative mechanisms are offered that will do the job of producing work which is filtered and selected before being published under a reputation-enhancing brand which signals authority? Currently 2,000

academic publishers (ranging from university presses, learned societies, independents large and small, through to the big conglomerates) do this through nearly 30,000 peer reviewed journals publishing around 1.5 million articles per year along with tens of thousands of academic monographs. But this scholarly output is only a drop in a vast ocean of information where 100,000 tweets are produced every minute, and attention is a much scarcer resource than information. When the attention economy is so overpopulated, it is easy for society to lose sight of the values and mechanisms that sustain authority in favour of the blunt and measurable traffic of information as commodity.

Knowledge may want to be free, but let us remember that not all knowledge is the same and that authoritative and lasting contributions in certain fields are sometimes produced under delicate, highly filtered conditions that enable and incentivise concentrated effort over time and a larger scale commitment to a disciplinary community. Let us not fall for an illusion like Immanuel Kant's light dove who in 'cleaving the air in her free flight, and feeling its resistance, might imagine that its flight would still be easier in empty space.'

Ziyad Marar *was born in Baghdad, Iraq in 1966. He lived in the Middle East until the age of 10 before moving to London. He holds a BSc in psychology (Exeter University), an MA in the philosophy and psychology of language, and did several years of postgraduate research in this field (University of London).*

He is Global Publishing Director of one the world's leading independent academic and professional publishers, SAGE. Having joined SAGE in 1989, Ziyad has worked across all aspects of publishing. During his career he has built international programmes in psychology and politics across journals, reference books, textbooks and online products. He was appointed Editorial Director in 1997, Deputy Managing Director in 2006, and took on his current global role in 2010, in which he has responsibility for the overall strategic direction of SAGE's publishing.

In recent years at SAGE Ziyad has also focused on supporting the Social Sciences more generally by working with many learned societies and key organizations such as the British Academy. Alongside extensive use of social media such as

*socialsciencespace (www.socialsciencespace.com) and socialsciencebites (www.
socialsciencebites.com), he has spoken (youtu.be/C73V8q8ZWtw) and written
(www.guardian.co.uk/higher-education-network/blog/2013/apr/08/social-science-
funding-us-senate) on this theme in various international contexts.*

*Ziyad is also the author of three books combining his interests in psychology and
philosophy,* The Happiness Paradox *(Reaction, 2003),* Deception *(Acumen,
2008) and most recently* Intimacy: Understanding the Subtle Power of
Human Connection *(Acumen, 2012). He lives in London with his wife and
three daughters.*

He can be followed on Twitter: @ZiyadMarar

Notes

1   www.huffingtonpost.com/don-tapscott/the-week-university-as-we_b_2566926.html

2   shirky.com/writings/broadcast_and_community.html

3   blogs.lse.ac.uk/impactofsocialsciences/2012/11/14/reflections-on-purpose

4   www.timeshighereducation.co.uk/422188.article

5   Karl Marx, 'A Contribution to the Critique of Political Economy', contained in the *Collected
Works of Karl Marx and Frederick Engels: Volume 29*, p. 270.

6   www.wm.edu/news/stories/2013/the-higgs-boson-looks-just-like-marc-sher-said-it-would.-
now-what123.php

7   Bertrand Russell, *Power: A New Social Analysis,* (London: Unwin Hyman 1938/1988), p.9.

8   en.wikipedia.org/wiki/Reflexivity_(social_theory)

9   en.wikipedia.org/wiki/Wicked_problem

10   Ibid.

11   www.nature.com/nature/journal/v495/n7442/full/495437a.html