

Downloads and Impact in Scholarly Book Publishing

Marcel Knöchelmann | Marcel.knochelmann.15@ucl.ac.uk

Department of Information Studies
University College London
119 Foster Court
Gower Street, London, WC1E 6BT

Published on *LePublikateur.de* in September 2018:

<https://www.lepublikateur.de/2018/09/25/downloads-impact-scholarly-book-publishing/>

It's hard to tell how often a scholarly book has been read. Neither digitalisation nor open access change that. Misrepresenting download numbers as impact or usage of books—paywalled or open access—isn't useful and can even be irresponsible, when this metric is used as a justification.

Performance Indicators in a Market-Driven Environment

The shift towards the digital is accompanied by a tendency to count things. Every digital process can easily be counted; a comparable process in a non-digital environment is often much harder to count, or at least more expensive. That's reason enough to take those statistics from the digital sphere and bestow upon them an often loosely connected argument: a metric is born.

Such a metric serves well in a neoliberal environment. Products are pushed into markets which return metrics that supposedly prove the products' performances. Considering that the arguments behind those metrics are fluid, those performance indicators are often used as justifications. Exactly this is what can be seen with download numbers of scholarly books, a metric that is poised to be trivial in many contexts.

Opening a Book in the Library and on the Platform

The problem is not the download number itself. In a contextualised representation, download numbers of a PDF/EPUB or visits of full HTML content can indicate potential interest in content and where the potential audience is located. In addition, those stats can be insightful as a comparative metric for a platform, for instance, regarding open access habitude in general. However, those metrics make no sense as a definitive statement about usage or impact of content.

This is where the argument behind a number becomes fluid. A download just can't reflect anything about the purpose of the download and what happened with the content after it has been downloaded. Sophisticated platforms can track usage: Amazon's Kindle can tell exactly who stayed how long on which position in the text (though, arguably, even this can't reflect much about the usage of the displayed text). If, however, someone downloaded a handful of items from OAPEN, how can this be taken as something else than a patron opening a few books in a library? The purposes for both downloading and opening may only have been to have a look at the ToC or the author's details. The difference is that on a digital platform this counts as impact and in a physical library it doesn't.

Open Access and the Deceptive Nature of Download Statistics

Even worse, those stats are less meaningful the poorer the design of the digital platform. If essential elements such as a ToC, author information, a long abstract, an exemplary passage of the text, etc., aren't displayed on the platform site, the potential reader must inevitably download the content to find out whether it's what she was looking for. This alone makes download stats quite useless as a comparative metric between sites.

The problem increases in an open access context. Downloads in a paywalled environment are arguably more representative of an intention, where overcoming a paywall or clicking through the Shibboleth authentication (or, indeed, visiting the pirate site) requires more will to action than clicking on the PDF button. While in both the paywalled and the open access environment the intention doesn't say anything about the motive, download statistics of free content are supposedly always higher than those of paywalled content. And who counts those myriads of circulating PDFs among fellow researchers? Are PDFs of open access publications shared more among researchers, or less (because you'd rather send around a link to the download button than attaching the PDF)? How can this uncertainty ever be represented in download stats?

Moreover, for the ordinary reader, downloads are more deceptive than anything. Take the accumulation of chapter downloads as an example ([seen here](#)¹): A book hasn't been downloaded 80,000 times if each of the book's 10 chapters has been downloaded 8,000 times. 8,000 downloads of a book would seem obvious. But 80,000 times just sells the platform better. Besides this platform-induced deception, there are bots downloading content or users accidentally clicking on download. COUNTER can be an improvement here, standardising the representation and verification of those statistics. But does the ordinary reader look up whether the platform's statistics are COUNTER-verified, or what COUNTER means in the first place?

Download Statistics as a New Impact Regime?

For those who are pushing the download metrics, the uselessness of these numbers isn't of much interest. And this is where the issue is becoming precarious. If it's not just a service provider playing around with numbers on the platform, but publishers and libraries judging

¹ Bookmetrix by SpringerNature: <http://www.bookmetrix.com/detail/book/044b3a55-b95d-422e-9130-978cf159ce83#downloads>.

the researchers' and their content's impact—indeed scholarly pursuit—based on such metrics, this, then, would be downloads as a new impact regime.

The word impact already seems to have become crucial for many processes in academia, especially scholarly publishing. Only the metrics behind the impact statements cannot fulfil the role of an argument for those statements. Impact, as 'the effective action of one thing or person upon another; the effect of such action; influence; impression' (OED²), would require at least a notable effect or at best a significant repercussion. Download numbers reveal none of this except for the action of downloading itself.

Open Access and its Justifications

Open access is in this light increasingly used to achieve more impact, and, the other way around, commercial open access approaches try to justify their open access costs through impact statements. This starts with the reasoning that the success of open access could be proved by off-campus download stats, where non-academic audiences supposedly downloaded scholarly titles. No one knows who is behind those downloads, so that the only proof of those stats is that online platforms have no geographic boundaries.

Moreover, a justification of expensive commercial open access approaches through misleading usage statements has a detrimental effect on educating researchers towards more openness. Open access is about ensuring equal access to knowledge for both readers and authors, and to rid the communication system of profit-seeking behaviour. Especially in the slowly-developing books environment, luring in authors or librarians to pay high BPCs (book processing charges) by promising more impact and usage is potentially playing into the assumption that open access created a new two-class environment: those who can afford to purchase additional impact and those who cannot. Yet, green is likewise a good

² Oxford English Dictionary: <http://www.oed.com/view/Entry/92036?rskey=a62khf&result=1#eid>.

solution for the transition, and the mentality change towards more openness should be based on more equality—not more competition.

Impact in the Context of the Humanities

Surely, open access brings higher visibility and potentially more citations and online mentions. But it's the value system behind the statements that is misleading and conceivably damaging, starting with the question, what is impact in the humanities? How can value of research in the humanities be expressed? Is there the need to force a metric in open access publishing? And most of all, with all those discussions about impact and value of scholarship in the humanities³, aren't we better served pronouncing the qualitative over the quantitative?

Altmetrics are quite new as well. And, like downloads, the availability of altmetrics statistics have only limited meaning. They are handy in the way that they provide an overview of the inclusion of the published material in other (online) content. But that doesn't equal impact. It's a well-known issue that referencing works in the humanities is often much more than a proof or a pointing towards further explanations. It's an interpretation, a discussion, an apologia, a defence, or any reference the value of which cannot be expressed in a +1. In such a context, how can download numbers be considered impact?

Researchers in humanities disciplines know this themselves and disregard of using downloads as a reference to their publications' impact. The more important aspect is that those involved in managing scholarly publications—librarians, publishers, academic decision makers—don't fall for a new regime of meaningless metrics.

³ For instance, C.P. Snow *The Two Cultures and the Scientific Revolution* (1959), Stanley Fish *Will the Humanities Save Us?* (2008), Frank Donoghue *Can the Humanities Survive the 21st Century?* (2010), Martha Nussbaum *Not for Profit: Why Democracy Needs the Humanities* (2010).

Qualitative Decision Making

The decision for publishing a work in the humanities, indeed for the scholarly pursuit in the first place, is entirely based on qualitative information. No meaningless metric should interfere with this process. Beware the metric regime that tells a future editor: ‘Well, having made that monograph on medieval poetry open access wasn’t a good idea according to the below-average download numbers. You better not include the next submissions in medieval literature in our costly open access programme.’ That’s neither how meaningful publishing programmes are crafted over time, nor what the reasoning behind open access is. That a submission is being published is much more than a decision based on gut feeling, and that it is made freely available is a service to fellow researchers and the public. It’s what the editor and her publishing house considers the right content for their programme, the right publishing mode for the content,⁴ and as representative for their brand.

Looking at download numbers may be an insightful bibliometric exercise. But it’s a number without much meaning for the content. Libraries, publishers, and researchers should refuse to be pulled into another impact regime. Metricised impact is neither at the heart of open access nor of publishing, and it’s toxic especially where scholarly pursuit is of qualitative, ideographic, or interpretative nature. Resist the meaningless metrics!

⁴ Besides policies and funding requirements interfering with this decision making.