bjoern.brembs.blog

7 functionalities the scholarly literature should have

Björn Brembs

Published August 1, 2017

Citation

Brembs, B. (2017, August 1). 7 functionalities the scholarly literature should have. *Bjoern.brembs.blog*. https://doi.org/10.59350/d441b-tsh82

Keywords

Science Politics, Functionality, Innovation, Literature, Scholarship

As a regular user of the scholarly literature since before the internet (I started reading primary scientific literature for a high-school project around 1989), I have closely followed its digitization. I find it rather frustrating that some of the most basic functionalities we have come to expect from virtually every digital object are still excluded from scholarly articles, making the literature much less useful than it could be. Some of these functionalities are more than 20 years old.

Nobody told Zuckerberg that they wanted Facebook. One would think that with profit margins sometimes exceeding 40% on billions in revenue, academic publishers would use some of that cash to provide their readers with at least a modicum of modern functionalities, without constant prodding. Alas, it appears as if these publishers have different priorities. They sue Sci-Hub, send take-down notices to Academia.edu for some papers, or they buy start-ups for reference management, preprint servers, laboratory notebook providers or companies providing altmetrics. It seems they don't really know what to do with all this cash burning holes in their pockets! I have some long overdue features publishers could instead invest the money they stole from us in (no particular order):

1. Accessibility

Today, most human readers with an internet connection can, after an often awkward and cumbersome process that sometimes involves several search engines and a variety of other tools read every digital scholarly article. However, in the age of Big Data, mining the scholarly literature for content is something so fundamental that it boggles the mind how publishers can get away with simply blocking this kind of research arbitrarily.

2. Smooth peer-review

Peer-review, as any similar social endeavor, involves humans with differing opinions, approaches and social skills. Today's version of formal peer-review of the scholarly literature has been around since around the 1950s. Given the time elapsed, one would think that the realization ought to have set in, that one shouldn't needlessly compound a process already fraught with trials and tribulations with cumbersome and clunky technical procedures. Today's common procedures run the risk of amplifying the chances of misunderstandings and simultaneously exacerbating anxiety and misbehavior of the people involved.

Those of us old enough to be able to read and write at this time, have enjoyed web-based message boards (or online forums) since at least 1994. Online commenting and annotation on web-based word processors such as, e.g. Google Documents have been around since about 2005. Compared to the current practice of a single text review and single text replies by the authors, these ancient, in web terms, technologies appear almost like magic. How can publishers in 2017 get away with service from before 1994? In particular since this idea, entered by Koen Hufkens received an award way back in 2012?

3. Social components

Social media technology started around 1999 with the development of the "web 2.0". However, our scholarly literature is still firmly stuck in web 1.0, despite the commendable efforts of some lone publishers to implement one of the earliest social functionalities, commenting. However, commenting is only one of many social technologies and one that may even better be implemented in a formal peer-review process at that. There is little reason we shouldn't be able to form scholarly online communities which share common interests – after all, scholarly societies have been around for centuries. These communities could use social functionalities to share articles, but also to track recommendations, citations, downloads, etc. as we do with any other digital object. Obviously, such a functionality only makes sense as a built-in property of the literature and not of some duplicated space where some users share some of the literature, à la RG *et al*. Such basic functionality would let the reader know who in their communities are reading which articles and which colleagues are publishing which of their results.

Remember the "customers who have bought this book also bought this one" recommendation from Amazon? That was around 1994, 23 years ago – and still no sign of our scholarly literature implementing this ancient feature as a native component.

4. Web-based data visualizations

These days, in our private lives, we routinely zoom through all kinds of map-like data, either with pinching fingers or the scroll-wheel. We rotate all kinds of 3D objects or dynamically adjust the graphs displaying the usage statistics of our personal sites. However, for the scholarly literature, almost exclusively, the one visualization that counts is the pixel-based, flat image that only displays what the authors want their readers to see. While some journals are demanding their authors deposit their data with the publication of their article, this is not to ensure proper visualization of said data. Instead, it becomes a tedious, manual process by which authors may pay lip-service to data accessibility in principle, without any added benefit to scholarship other than the theoretical ability of a select few experts to potentially have a second look at the data (also that second look being cumbersome and manual). As if to add insult to injury, nobody seems to care about the software we write to transform the bits and bytes of the raw data into the flat, pixel-based images.

5. Hyperlinks

The first public demonstration of hyperlinks was also the first time a computer mouse was demonstrated. It was in 1968, in what is today called "the mother of all demos". In the almost 50 years since then, we haven't managed even to properly implement hyperlinks into the scholarly literature. For example, try and click on a very common sentence, e.g. "the experiments were performed as previously described". In essentially every single case today, nothing happens, while in the demo in 1968, it would have taken the reader to a document describing the experiments.

If today's reader is lucky, there is a reference behind the sentence that is clickable. However, in the majority of cases, it will just lead the reader to the place in the reference list where the full reference is listed. Whether this reference is clickable remains a hit or miss. In the affirmative case, however, the click will still not lead to a description of the experiments, but to a paper (or

a paywall). If that paper is accessible, today's reader may again be lucky and find the particular experiment buried somewhere in the Materials and Methods section, or, if less lucky, only some components with further references.

Imagine what would happen to an online store that would treat customers like that if they wanted more detailed descriptions of the merchandise. As if it wasn't already clear before, this comparison should make it quite obvious what academic publishers think of their readers.

6. Semantic web technology

Coincidentally, proper hyperlinking of our literature with standard technology from about a decade ago, would also provide every article with an automatic list of citing articles (i.e., pingbacks/trackbacks) and allow deep citations (aka. anchors) to text, data and code. This would simultaneously allow us to implement a citation ontology to specify what kind of citation we are using, with myriads potential use cases, all benefiting scholarship.

Speaking of ontologies: semantic mark-up of all our scholarly articles would greatly benefit all kinds of scholarship, be it systematic reviews, content mining or simple literature searches, etc.

7. Single click submission

Once I click "publish" on this blog post, it is made public, without any further ado. Yet, despite institutional logins, ORCID, etc. the vast majority of publishers still require us to fill in forms about each author, copy-and-paste titles, abstracts and other information as if we just put a stack of paper sheets into an envelope, rather than a document which already contained all this information.

I'm sure there are more candidates for obvious functionalities every user of the scholarly literature would like to have, but I'll leave it at that for now. At least no academic publisher can claim any more, they didn't know what their users wanted.