Fast and made to last: Academic blogs look to ensure long-term accessibility and stability of content.

LSE blogs.lse.ac.uk/impactofsocialsciences/2015/04/30/fast-and-made-to-last-academic-blogging/

Academic blogging has distinct advantages over traditional forms of scholarly communication but questions on their lasting preservation still remain to be seen. Who makes sure academic blog content stays online in the long term? Who guarantees that links to the post remains the same? Who ensures that the text will not be modified later on? **Christof Schöch** *argues these are issues that need to be resolved if blogs are to be reliable, trusted, citeable resources that receive academic recognition.*

More and more researchers understand how useful it can be for an academic field to have a lively ecosystem of research blogs. For example, researchers can use blogs to publish glimpses into ongoing work, ahead of more formal journal article publications. In this way, the community can be informed of the latest developments, without the overhead and delays of peer-review and other editorial processes. A conversation can start early on, providing feedback and driving the development of further research. One post can spark comments or companion posts on related topics. (An example for this type of blog posts can be found in computational literary text analysis, where blogs such as those by Ted Underwood, Matthew Jockers, Ben Schmidt and others provide the most up-to-date picture of the field.) This specific way of using research blogs for fast publication means that blogs can fulfil a function similar to that of talks at workshops or other, smaller and informal academic events. However, while face-to-face discussions can be very fruitful, they are bound to a certain time and space, and they are fleeting phenomena.

Image credit: Anonymous Account (Flickr, CC BY)

The advantage of blogs compared with such talks is that here, discussions can happen across geographical and temporal borders, and that they stay visible online in comments or companion posts. But aren't blog posts, ultimately, almost as fleeting as a talk at a workshop? Who makes sure the content stays online not just today and

tomorrow, but in the long term? Who guarantees that the link to the post remains the same? Who ensures that the text will not be modified later on? These are issues that need to be resolved if blogs are to be reliable, trusted, citeable resources and receive academic recognition even in the absence of traditional pre-publication peer-review. (Another source of academic recognition comes, of course, from post-publication peer review, as it is practised, in the field of Digital Humanities, by actors such as DHNow or the Journal of Digital Humanities, who give visibility to selected content once it has been published elsewhere.) The research blogging platform hypotheses.org has understood this early on. This fact is undoubtedly a factor in the success of the platform, which is run by the French initiative OpenEdition and currently hosts 1006 (and counting) research blogs in French, Spanish, Portuguese, German and English coming from the Humanities and Social Sciences. (Disclaimer: The author is member of the editorial team of the German branch of hypotheses.org.)

So, what does hypotheses.org currently do to ensure the long-term accessibility, citeability and stability of the content published on the platform, and what can still be improved?

- Whenever a blog has shown some degree of constant activity over a certain period of time, and if it is not designed from the outset to be a temporary blog (linked to a conference or a teaching activity, for instance), it receives an ISSN number from the relevant national library (such as the Bibliothèque nationale de France or the Deutsche Nationalbibliothek) or the international ISSN service. This makes the blog a formally registered periodical publication and helps with clearly identifying the blog. This does not mean, however, that the blogs are automatically added to the online-content archiving activities of the national libraries.

- OpenEdition, the organisation behind hypotheses.org is a well-established institution also running the platform revues.org (for online journals) and Open Edition Books (for digital monographs). It will make sure the blogs and their content will remain online in the long run. There is no formal or legal guarantee for this, however, and no automatic archiving routine with some other institution or research content repository is currently in place.

- Each individual blog post receives a unique numerical identifier used to construct its URL. These are purposefully not speaking URLs, because those tend to get very long if they are to be unique. However, it is important to keep them short because that makes it much easier to cite them in more formal, printed publications (an aspect, again, related to academic recognition).

- The URLs do not change, as neither the base URL of the blog nor the identification number of a blog post can be changed. However, they are not technically speaking permalinks or PURLS, and hypotheses.org does not currently provide DOIs (Digital Object Identifiers) for individual blog posts.

- As for the actual content itself, there are currently no technical barriers which would prevent authors from changing the content of a blog post after it has been published. Although WordPress, the blogging system used by hypotheses.org, stores a record of revisions to any blog post, these revisions are not accessible to readers of the blogs. There is only an implicit understanding in the community that any modification of a blog post which goes beyond fixing a typo or a link, should be made transparent to the readers in some way.

So, in a nutshell: hypotheses.org is doing quite a few things already to make blog content published on the platform easily and clearly identifiable and to make sure it stays online in the medium term. However, blog content may be changed and no repository (integrated or separate from hypotheses.org) records versions of blog posts. It would probably be a good start to add some information about the "best practice" of not changing blog content once it is publidhed to the existing guidelines made available by hypotheses.org. A bigger step forward would be to connect hypotheses.org with Zenodo, the European research repository associated with OpenAire and CERN, or with HAL, the French open archive. Stable versions of blog posts could be automatically archived there and given a DOI in the process. Some practical and legal issues would probably need to be resolved for this to happen, but it would be a step forward towards guaranteeing the stability and availability of the content published on the hypotheses.org blogs. In turn, this would help making research blogs a quick and flexible as well as a reliable and well-recognized way of fostering academic discussion online. Blogs would be both fast and made to last.

Note: This article gives the views of the author, and not the position of the Impact of Social Sciences blog, nor of the London School of Economics. Please read our comments policy before posting.

About the Author

Christof Schöch *is a research associate at the Chair for Literary Computing, University of Würzburg, Germany, where he leads the junior research group on Computational Literary Genre Stylistics. In his research and teaching, he combines French and Spanish literary studies and Digital Humanities. He writes about his research on a blog called "The Dragonfly's Gaze". You can find out more about him on his personal website.*