

EASE-Forum Digest: September to December 2016

You can join the forum by sending the one-line message "subscribe ease-forum" (without the quotation marks) to majordomo@helsinki.fi. Send in plain text, not HTML. Details at www.ease.org.uk/node/589.

Crawlers

The application form for inclusion in the Directory of Open Access Journals (DOAJ) asks journals if they allow anyone to crawl their full-text. Arjan Polderman wanted to know what was meant by text crawling. Andrew Davis explained that a crawler could be used by readers to automatically find and open databases, eg containing journal articles, search the content and supply an ordered list of its findings. He uses a crawler to find examples of particular enzymes active on specified substrates. Duncan Nicholas said that for inclusion in the DOAJ "The journal platform would need to offer a standard application programming interface (API) and standard data representations to facilitate this text-mining. CrossRef have provided a service for publishers to standardise the ways in which they can provide this access." Further details could be found at <http://tdmsupport.crossref.org/>, while Liz Wager also suggested looking at <https://www.springer.com/gp/rights-permissions/springer-s-text-and-data-mining-policy/29056>

However, a crawler as a computer program that systematically browses the World Wide Web to create an index of data is a new meaning of a word that describes a servile flatterer, a baby who has not yet learned to walk, or a thing that moves at a slow pace such as a heavy vehicle on tracks used in building construction.

PubMed's XML requirement

Sylwia Ufnalska had heard from a friend's friend that PubMed is now requiring journals to submit full articles in XML if they wish to continue to be listed after this year, which could increase costs for journals that do not use XML. Carmel Williams of the *Health and Human Rights Journal* responded that she was the friend's friend and her journal had to provide over 50 papers in XML format to meet PubMed's "technical assessment," which would continue to be a future requirement. This was a burden for a small journal but PubMed had sent them a list of vendors who undertook conversions from PDFs/InDesign or Word into XML. She had used a company in India which she could recommend to anyone who would like to contact her (williams@hsph.harvard.edu).

Andrew Davis wondered if recent versions of Adobe Acrobat or Word might not be used for the conversion to XML but Aleksandra Golebiowska thought this would not do as PubMed have their own XML format. In her experience dedicated programs were required as putting in a wrong code for a character had resulted in some of her files failing to be validated.

What is meant by the author contributions "substantive translation" and "editing"?

Contributor Roles Taxonomy (CRediT, <http://casrai.org/credit>) aims to standardise the description of authors' contributions to manuscripts to allow this metadata to be more easily retrieved through databases. Valerie Matarese was concerned that definitions used by CRediT conflict with those used by language professionals. Valerie queried

1. What was meant by the "substantive translation" contribution under "Writing — Original Draft"?
2. Whether the word "editing," which she considered to be an activity of non-authors, in the heading "Writing — Review & Editing" would be better replaced by the word "revision," an author activity?

Sylwia and Ana Marusic both agreed that authors revise rather than edit. Sylwia further supported this view by pointing out the text under the heading referred to "critical review, commentary or revision" and did not mention, "editing." Ana added that a number of studies undertaken by her group^{1,2} had shown that open questions to authors about their contributions resulted in more accurate answers than categorised questions. Valerie thought the CRediT working group was aware of Ana's research, as it was cited in their article explaining their aims in *Learned Publishing*,³ but seemed to hold the opposite opinion. However, no representative of the language profession had been included in their working group, even though their article stated, "CRediT brings together a diverse set of stakeholders wanting to better understand and communicate the different kinds of contributor roles in research outputs." Mary Ellen Kerans also thought questionnaires where authors only wrote ticks against a list were not working, for example, she had seen absurd article end notes generated from such questionnaires stating animals had not been harmed in studies where no animals had been used.

Mary Ellen thought "substantive translation" referred to a collaborative process between an author and translator that takes place before publication when no original language version had been published.⁴ The translator would only qualify under CRediT's "Writing — Original Draft" if the author's reading knowledge of the language of translation were insufficient, when the translator needed to sign the contributor statement and a footnote should be included to explain who wrote the unpublished manuscript.

As for "editing," Mary Ellen pointed out that the word is commonly used to describe what authors do when they polish their final draft. Therefore, she considered the word in the "Writing — Review & Editing" heading to be apt. However, she believed the word truly expressed the difference between "just" substantive editing and authoring and when, to quote Mary Ellen, "An editor works mainly with what the author's already put on the page." In this case, "An author works with more — generates further ideas, remembers forgotten ideas and experiences, and even has changes of opinion during what might have initially been hoped was a polish/edit of a last draft."⁵

In the meantime, Michael Altus had Googled “substantive translation.” He was unable to find a definition but the PLoS journals, which had adopted the CRediT guidelines, included the following under the “Writing – Original Draft Preparations” heading: “Creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation) (<http://journals.plos.org/plosone/s/authorship>). Indeed, we do not know if CRediT’s working group is aware of the definition offered by Mary Ellen. Michael suggested the editor of PLoS might be able to explain what “substantive translation” means for that journal.

Is placing a manuscript on an institutional website a prior publication?

Ivana Stetinova posed this question. A journal editor had accepted and published an article with a DOI as an online-first. The editor subsequently Goggled the article and found it on the web page of the author’s institute with a note that the paper could be cited through the website. The author had confirmed on submission that the manuscript had been submitted solely to the journal and was not published, in press, or submitted elsewhere. On being questioned, the author replied that the research’s funder required studies to be placed in the electronic journal “Science paper online” before publication in any other journal. The author argued that as “Science paper online” was just a communication platform and authors retained the copyright there could not be a problem. The author seemed to confuse copyright with prior publication. As to whether there was a problem, Andrew Davis considered this depended on if the journal considered a posting on an institutional website or preprint server to be a publication. COPE seemed to have no clear guidelines on the point but he thought as there was no active distribution of copies of the article the posting could be equated with one on *arXiv* and was not a prior publication. If the journal was unhappy with this conclusion it should state on its website that placing a manuscript on an institutional website would be considered a prior publication.

Elise Langdon-Neuner (compiler)
a.a.neuner@gmail.com

Contributors

Arjan Polderman: A.K.S.Polderman@pw.nl
Sylwia Ufnalska: sylwia.ufnalska@gmail.com
Valerie Matarese: vmatarese@uptoit.org
Stetinova Ivana: stetinova@ueb.cas.cz

References

- 1 Marusic A, Bates T, Anic A, Marusic M. How the structure of contribution disclosure statements affects validity of authorship: a randomized study in a general medical journal. *Current Medical Research and Opinion* 2006;22(6):1035-44. Available at <https://www.ncbi.nlm.nih.gov/pubmed/16862642>

- 2 Malicki M, Jeroncic A, Marusic M, Marusic A. Why do you think you should be the author on this manuscript? Analysis of open-ended responses of authors in a general medical journal. *BMC Medical Research Methodology* 2012;20:12-189. Available at <https://www.ncbi.nlm.nih.gov/pubmed/23256648>
- 3 Brand A, Allan L, Altman M *et al*. Beyond Authorship: attribution, contribution, collaboration, and credit. *Learned Publishing*. 2015;28:151-155
- 4 Shashok K, Kerans ME (2001) Translating the unedited science manuscript: Who fixes what shortcomings? In: Chabás J, Cases M, Gaser R, editors. Proceedings: First International Conference on Specialized Translation. Barcelona: Universitat Pompeu Fabra, p101-4.
- 5 Kerans ME. Writing process research: implications for manuscript support for academic authors. In: Matarese V, editor. Supporting research writing: roles and challenges in multilingual settings. Oxford: Chandos; 2013. p39-54.

EASE Conference 2018

Bucharest, Romania

No firm date yet for your diary,
but probably in June.

Ideas for topics welcome:

- overall theme
- plenary lectures
- parallel sessions
- workshops
- any other ideas?

Please send all ideas and suggestions to Joan Marsh
joan.marsh@lancet.com

New EASE members

EASE would like to welcome the following new members:

- Madalina Georgescu, Romania
- Karolina Kalisz, UK
- Paul Beverley, UK
- Anna Sawicka, France
- Sherryl Sundell, Germany
- Petal Smart, Trinidad and Tobago
- Vasuprada Iyengar, Germany
- Joanna Schuurman, Switzerland
- Zafer Kocak, Turkey