Review of "Research Articles in Simplified HTML: a Web-first format for HTML-based scholarly articles"

Cite this review as:

Soiland-Reyes S (2017): Peer Review #3 of “Research Articles in Simplified HTML: a Web-first format for HTML-based scholarly articles (v0.1)”. PeerJ Computer Science https://doi.org/10.7287/peerj-cs.132v0.1/reviews/3

- Authors: Silvio Peroni, Francesco Osborne, Angelo Di Iorio, Andrea Giovanni Nuzzolese, Francesco Poggi, Fabio Vitali, Enrico Motta
- Title: Research Articles in Simplified HTML: a Web-first format for HTML-based scholarly articles aka https://doi.org/10.7717/peerj-cs.132
- Submitted to: PeerJ Computer Science (2016-10-10)
- Reviewed by: Stian Soiland-Reyes (3/3) (2017-01-12)
- Submitted version: https://w3id.org/people/essepuntato/papers/rash-peerj2016/2016-10-08.html aka https://doi.org/10.7287/peerj.preprints.2513v1
- Published version: https://doi.org/10.7717/peerj-cs.132
- Dataset: https://dx.doi.org/10.6084/m9.figshare.3980463
- Editor decision: Minor revisions

Comments for the Author

This article presents RASH, an HTML-based format for authoring, exchanging and publishing academic articles, arguing that this allows a "Web-first" approach to authoring with a focus on content; but with facilities for semantic annotations. The associated RASH software allows conversion to more traditional article styles (for PDF), as well as conversion from a traditional word processor to RASH HTML.

While the authors argue that we should aim for a "Web first" publishing model with no conversion to traditional PDFs, as championed by the perhaps more visionary linkedresearch.org movement, here it is proposed that RASH gives a pragmatic approach that requires smaller adaptation to existing co-authoring and publishing workflows.

The authors has performed a kind of usability survey for RASH users at two consequent workshop, which gives validity to the claims of its purpose, but also (as recognized by the authors) highlights the current gap in tooling and skills to produce the underlying HTML and RDF annotations.

I think RASH can be seen as an important element of modernizing academic publishing; and while it can be argued that a restricted HTML template like RASH can limit academic authors from publishing articles augmented with state of the art web technology (for instance for interactive data rendering), this model is also a stepping-stone with a stronger focus on portability and longevity that lowers the barriers to get existing publishers on board.

Basic reporting

Detailed comments (including grammar suggestions) are available as hypothes.is annotations on the RASH version of the article: https://via.hypothes.is/https://essepuntato.github.io/papers/rash-peerj2016.html (Expand on the right, see inline comments)

The language of the article is OK, but I’m afraid it needs some work several places to improve clarity, e.g. by rephrasing or simplification.
Some of the data (the CSV files) has been shared on Figshare and cited as such, but I am missing the raw data of the extracted RDF annotations as well as the scripts used for extraction.

The HTML file of the article in RASH format has for some reason not been submitted as an additional file (only cited by URL) - perhaps this was not supported by PeerJ's submission system?

The RASH framework and associated software is referenced by GitHub URLs, but without using versioning. For archival purposes and future availability I would appreciate a Zenodo or Figshare archive of a tagged/version of the software, cited using DOI.

**Experimental design**

No Comments.

The article describes well the rationale and methodology.

**Validity of the findings**

The survey data is robust, but perhaps of a bit small sample size to be statistically sound. This is however helped by the fact that the survey was run over two consequent years.

The paper makes several unfounded claims using words like "guarantees" - I believe this is more of an English language/grammar issue than actual claims.


---

**Re-review of "Research Articles in Simplified HTML: a Web-first format for HTML-based scholarly articles"**

Cite this review as:

Soiland-Reyes S (2017): Peer Review #3 of "Research Articles in Simplified HTML: a Web-first format for HTML-based scholarly articles (v0.2)". PeerJ Computer Science https://doi.org/10.7287/peerj-cs.132v0.2/reviews/3

- Authors: Silvio Peroni, Francesco Osborne, Angelo Di Iorio, Andrea Giovanni Nuzzolese, Francesco Poggi, Fabio Vitali, Enrico Motta
- Title: Research Articles in Simplified HTML: a Web-first format for HTML-based scholarly articles aka https://doi.org/10.7717/peerj-cs.132
- Submitted to: PeerJ Computer Science (2017-07-06)
- Reviewed by: Stian Soiland-Reyes (3/3) (2017-08-06)
- Submitted version: https://w3id.org/people/essepuntato/papers/rash-peerj2016/2017-07-06.html aka https://doi.org/10.7287/peerj.preprints.2513v2
  - Revision of: https://w3id.org/people/essepuntato/papers/rash-peerj2016/2016-10-08.html aka https://doi.org/10.7287/peerj.preprints.2513v1
  - Annotated by reviewer #3: https://via.hypothes.is/https://w3id.org/people/essepuntato/papers/rash-peerj2016/2017-07-06.html
- Dataset: https://doi.org/10.6084/m9.figshare.3980463.v5
- Editor decision: Accept
- Accepted version: https://doi.org/10.7717/peerj-cs.132

**Comments for the Author**

I am Stian Soiland-Reyes http://orcid.org/0000-0001-9842-9718 and believe in open reviews.

This review is licensed under a Creative Commons Attribution 4.0 International License
The authors present the RASH framework, a subset of XHTML for academic publishing, along with software tools for its validation and conversion. The paper review the state of art in academic HTML publishing, motivate and detail the design of the framework, and evaluate its uptake and future challenges.

While personally I would have welcomed a more visionary/revolutionary approach for changing academic publishing for the Web, the authors take a more conservative approach with emphasis on pragmatic tooling to support existing authoring workflows (e.g. support LaTeX and MS Word). From this, RASH can provide a valuable stepping stone for more structured and accessible Web publication workflows for academic publishing.

I think this is a solid article that presents an important contribution to the further development of web-based scholarly communication.

I Recommend this article as "Accept" - although I have left some annotations in https://via.hypothes.is/https://essepuntato.github.io/papers/rash-peerj2016/2017-07-06.html#rash-eval (Click "<" and "Show Annotations") that I hope the authors will consider (along with this review) if a revision nevertheless take place.

Basic reporting

The paper have clear language, significantly improved since first review. The dataset is augmented with extra material, and referenced properly from Figshare with https://doi.org/10.6084/m9.figshare.3980463.v5

Argumentation is well-structured and founded, although a couple of citations or examples are missing, e.g. claim that HTML allow ambiguous structures, or the novel (and unnecessary) use of RDF/XML in a script tag.

See https://via.hypothes.is/https://essepuntato.github.io/papers/rash-peerj2016/2017-07-06.html#rash-eval for my detailed review per section of this version.

Experimental design

The paper describes well the motivation and design of the RASH framework, while also giving an extensive and up to date review of comparative technologies and approaches. The paper explains also challenges and peculiarities encountered in its implementation.

RASH is a well-designed subset of HTML that emphasizes document structure and semantic annotations. I think it could also be argued that unlike "any HTML5" this design also improves longevity for articles published in RASH HTML.

My only slight concerns is the extension of WAI-ARIA roles (e.g. "doc-endnotes"), which I could not find any citations for being allowed (or not) within HTML5; as well as the novel use of RDF/XML in a HTML script tag.

Validity of the findings

The survey part provides valuable insight into the uptake potential of RASH-like technology - although this should be taken with a grain of salt as the relative low number of survey participants means the data is (as the authors point out) NOT statistically significant. The paper do however provide a good qualitative analysis of the findings, which warrants their inclusion.

The authors provide well-reasoned conclusions. While my previous review identified some speculative language, this have now been improved.