

Archiving, but not really

Egon Willighagen 

Published August 6, 2025

Citation

Willighagen, E. (2025). Archiving, but not really. In *chem-bla-ics*. chem-bla-ics. <https://doi.org/10.59350/vwd81-p8z85>

Keywords

Publishing

Abstract

Mike Taylor wrote up a post about the various things a journal article is doing, the first being a scientific report. We put a lot of money in establishing a scientific track record. In the past 30 years how we publish our research and how we archive it has changed significantly. If you read my blog more often, you know I have been critical of the performance of many publishers.

Copyright

Copyright © Egon Willighagen 2025. Distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

chem-bla-ics

So, this is **really bad**! The archived version is not really useful. As a human I already struggle to read the degraded image, let alone an algorithm.

Does that matter? Yes, projects like the awesome [Pathway Figure OCR](https://doi.org/10.1186/s13059-020-02181-2) (see doi:[10.1186/s13059-020-02181-2](https://doi.org/10.1186/s13059-020-02181-2)) depend on images to be FAIR enough to extract information. (Side note: yes, these images should be vector graphics, but commercial publishers decided about twenty years ago that they could not care enough.)

At this moment, I do not know where the information is lost. Maybe PubMed Central is storing the images in a low resolution. Maybe the publisher provides PMC with a low resolution image. But to me, this must be solved as soon as possible. This is utterly unacceptable.

I wonder what the authors of the article (doi:[10.1186/s13287-025-04166-z](https://doi.org/10.1186/s13287-025-04166-z)) I took as example think of this.