

Archiving blogs

Egon Willighagen 

Published December 27, 2024

Citation

Willighagen, E. (2024, December 27). Archiving blogs. *Chem-bla-ics*. <https://doi.org/10.59350/vjvdy-6p110>

Keywords

Blog, Openlab

Copyright

Copyright © Egon Willighagen 2024. Distributed under the terms of the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

chem-bla-ics

Blogs come and go. Sometimes they move from one location to another. However, blogs have not been systematically archived, perhaps for work by efforts by OpenLaboraty. Bora Zivkovic gave in 2012 [a good overview](#) , to which Paul Raeburn [replied](#): *“If you weren’t blogging in the mid-2000s, when all the science bloggers knew and blogrolled each other, you’ve already missed the golden age.”* I think blogging is as strong as ever, but a lot of blogs have become more like columns in bigger media. Archiving of blog had not been done systematically, tho some posts made it into print, for example in [the Open Laboratory](#) series. Some copies made it into libraries, e.g. [2006](#), [2010](#), and [2012](#).

The two posts from the first paragraph from the [blogs.scientificamerican.com](#) provide a good example of the problem: bitrot. The Internet Archive has always been useful for archiving webpages and has been useful for archiving blogs too. But I do not believe it has been used systematically either, but at least it helped recover the above two pages.

So, when I discussed [the blog of Rich Apodaca earlier this month](#), the question came up if we could archive his blog. Beside his [personal coverage](#) of his cancer, his blog also covers a good bit of open science cheminformatics of the zeros and 10s.

Rogue Scholar

This is where [Rogue Scholar](#) comes in. [Martin Fenner](#) took up my question and started archiving Rich’ blog, resulting in [this ‘community’](#) collecting the blog posts. This is what an archive page for a single blog post looks like:

Published November 23, 2022 | Version v1 | <https://doi.org/10.59350/ecrpn-2gs23>

Balsa Reference Implementation

Apodaca, Richard L. 

SMILES is a compact molecular serialization format used widely in cheminformatics and computational chemistry. Even so, the published documentation incompletely describes SMILES, making the implementation of software impossible without some degree of improvisation. Balsa is a fully-specified language subset created to solve this problem. To bridge theory and practice, work on a Balsa reference implementation is well underway.

Additional details

Identifiers

UUID

26ab73cd-d869-4819-8902-76ebfd7adf5a

GUID<https://depth-first.com/articles/2022/11/23/balsa-reference-implementation/>**URL**<https://depth-first.com/articles/2022/11/23/balsa-reference-implementation>

Dates

Issued

2022-11-23T16:00:00

Updated

2022-11-23T16:00:00

What this archive now has is DOIs for each blog post, archived metadata that will also propagate via DataCite, etc. It does not have PDFs or other copies of the full blog posts yet. There are more than 900 blog posts to create PDFs for. Anyone [has an idea?](#)

I will post later this year about formally/semantically linking blogs citing other blogs using DOIs for blog posts, for example from Rogue Scholar. Any probably throw in [some use of the Citation Typing Ontology](#).

chem-bla-ics

Anyway, I can recommend everyone to get their blog lists on Rogue Scholar, for the DOIs and for the automatic archiving.