chem-bla-ics

Slow publishing innovation: SMILES in ACS journals

Egon Willighagen

Published February 21, 2014

Citation

Willighagen, E. (n.d.). In *chem-bla-ics*. chem-bla-ics. https://doi.org/10.59350/j5a8t-j7z67

Keywords

Publishing, Smiles, Acs

Copyright

Copyright © Egon Willighagen 2014. 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.

Elsevier is not the only publisher with a large innovation inertia. In fact, I think many large organizations do, particularly if there are too many interdependencies, causing too long lines. Greg Laundrum made me aware that one American Chemical Society journal is now going to encourage (not require) machine readable forms of chemical structures to be included in their flagship. The reasoning by Gilson *et al.* is balanced. It is also 15 years too late. This question was relevant at the end of the last century. The technologies were already more advanced than what will now be adopted. 15 years!!! Seriously, that's close to the time it takes to bring a new drug on the market!

Look at what they suggest and think about it. Include SMILES strings for structures in the paper. I very much welcome this, of course, despite I am not a big fan of SMILES at all. They could have said something about OpenSMILES too, which is more precise. They do say something about the InChI and InChIKey, but not that the SMILES string can more precisely reflect the drawing. I wonder why they don't go for a format that can actually capture the image, like CML or a MDL molfile. Then again, a SMILES copy/pastes so nicely. Talking about slow innovation. There is zero technical reason you could not copy/paste a MDL molfile into a spreadsheet (and you can with many tools, in fact...)

Now, I still have tons of questions. What tool will be used to validate the correctness and absence of ambiguity before the publication? Will the SMILES strings be validated at all? And at what level? Will it have to be compatible with particular tools? Does it have to be compatible with OpenSMILES? Under what license will these SMILES be available (can we data mine DOI-SMILES links and openly share them)? What was the reasoning for finally adopting this? Will the journal also accept submission where both SMILES and other formats are provided? Will they accept or deny SMARTS strings (e.g. for Markush structures)?

All in all, I second the others, and am happy to see this step. I do hope they do not stop here and wait again 15 years for another step. In fact, they ask for input on jmc@jmedchem.acs.org. That is double promising!