New paper: A template wizard for the cocreation of machine-readable data-reporting to harmonize the evaluation of (nano)materials

Published May 27, 2024

Citation
Willighagen, E. (2024, May 27). New paper: A template wizard for the cocreation of machine-readable data-reporting to harmonize the evaluation of (nano)materials. *Chem-bla-ics*. https://doi.org/10.59350/vfvwq-s0v13

Keywords
Rdf, Opentox, Fair

Abstract
I was about to call this blog post From spreadsheets to RDF, after the post last week. But then I decided to just use the pattern I typically use. Why I wanted to use that shorter term in the first place was that one of the thing I like about the AMBIT software (of OpenTox and eNanoMapper fame) is its RDF support (see doi:10.1186/1756-0500-4-487). But RDF, ontologies, those are hard things.

I was about to call this blog post *From spreadsheets to RDF*, after the post last week. But then I decided to just use the pattern I typically use. Why I wanted to use that shorter term in the first place was that one of the thing I like about the AMBIT software (of OpenTox and eNanoMapper fame) is its RDF support (see doi:10.1186/1756-0500-4-487). But RDF, ontologies, those are hard things. And unlike mathematics, we do not have simple objects like integer numbers or simple operators. Well, I think we do, and we talk about them. But there is no obligatory education. Just like any biologist needs to know what *1 + 2* means, I think any biologist needs basic knowledge about how knowledge graphs work. But sometimes feels like a taboo, like cursing in the life sciences church.

So, there we are. This is where spreadsheets come in. If done well, they combine aspects of knowledge graphs with usability and can even cover a good bit of the learnability. This is what is described in this new paper about templates in the EU NanoSafety Cluster: *A template wizard for the cocreation of machine-readable data-reporting to harmonize the evaluation of (nano)materials* (doi:10.1038/s41596-024-00993-1).

The learnability comes in with the spreadsheet templates (“this is how we did it”) and a “wizard” around it guides the user with the selection of a template but also can provide feedback on the template. The technical term for that is “validator”, but it can be tought of as a spelling checker. Computers are good at finding contradictions (the lack of a pattern), though less good at ranking the alternatives (which is the cause of hallucinations in AI approaches).

And to return to the RDF, software like AMBIT can read these templates, use the semantics linked to the template, and make the FAIR static spreadsheets (good for archiving on Zenodo!) available as FAIR interactive data (good for exploration and machine learning), and as RDF (good for data integration).

Congrats to Nina and the various EU NanoSafety Cluster projects!