

New paper: From papers to RDF-based integration of physicochemical data and adverse outcome pathways for nanomaterials

Egon Willighagen 

Published May 20, 2024

Citation

Willighagen, E. (2024, May 20). New paper: From papers to RDF-based integration of physicochemical data and adverse outcome pathways for nanomaterials. *Chem-bla-ics*. <https://doi.org/10.59350/jdj8r-h6187>

Keywords

Fair, Rdf, Erm

Abstract

Making something FAIR is hard, particularly when you do more than making something findable. We've seen before that making something usefully findable requires deep indexing, and already that continues to be difficult, because we are not seeing it enough. So, when I thought convert a paper led by Hoet's lab in Leuven into machine-actionable RDF to make it FAIR, I gravely underestimated the amount of work.

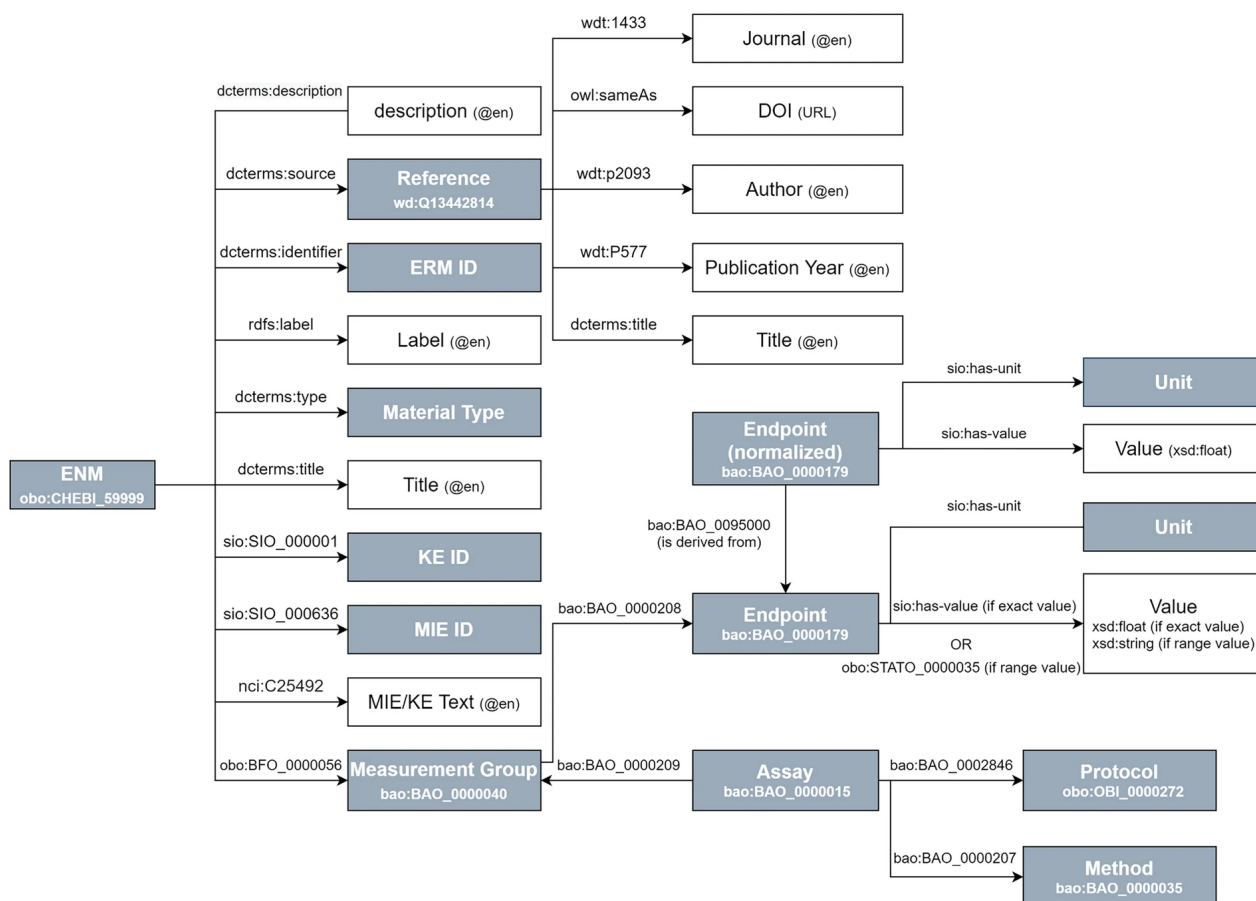
Copyright

Copyright © Egon Willighagen 2024. 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.

Making something FAIR is hard, particularly when you do more than making something findable. We've seen before that making something usefully findable [requires deep indexing](#), and already that continues to be difficult, because we are not seeing it enough. So, when I thought convert a [paper led by Hoet's lab in Leuven](#) into machine-actionable RDF to make it FAIR, I gravely underestimated the amount of work. [Jeaphianne](#) et al. did an awesome job on this work (doi: [10.1186/s13321-024-00833-0](https://doi.org/10.1186/s13321-024-00833-0)).

The idea was simple: write up which nanomaterial (type) activates which molecular initiating event. It would simply annotate each material with a unique identifier to link it to databases like [eNanoMapper](#) and [NanoCommons](#) and it would use unique identifiers for the [Adverse Outcome Pathway](#) (AOP) key events. As such, it would make a direct link in the growing linked open data cloud between the AOPs and the nanomaterial databases.

Unfortunately, it was quickly discovered that actually reusing this new datasets requires rich annotation (metadata!) of the materials and the materials from the source paper were not yet in material databases. And then the cumbersome start was started, resulting in a very rich data model describing the key events, the materials, the assays used, and the original papers themselves:



But the work has not finished yet. The paper assigned [ERM identifiers](#) to all included materials, and now these need to be added to new [ERM Identifier Database](#) under development.