Answering scientific questions with linked European nanosafety data

Egon Willighagen[1], Micha Rautenberg[2], Linda Rieswijk[1], Friederieke Ehrhart[1], Jiakang Chang[3], Georgios Drakakis[4], Penny Nymark[5], Pekka Kohonen[5], Gareth Own[3], Haralambos Sarimveis[4], Christoph Helma[2], and Nina Jeliazkova[6]

[1] Maastricht University, Maastricht, NL
[2] in silico toxicology Gmbh, Freiburg, DE
[3] EMBL-EBI, Hinxton, UK
[4] NTUA, Athens, GR
[5] Misvik AB, Turku, FI
[6] IdeaConsult Ltd., Sofia, BG

Nanomaterials are increasingly used in healthcare and consumer products. The safety of these materials seems adequate, but the European community seeks to back this up with experimental research data. Ideally, read across and predictive toxicology approaches can then be used to answer questions if a class of metal oxides is genotoxic or not. If successful, this will replace animal testing in bringing new nanomaterials to the market.

The eNanoMapper project (http://enanomapper.net/) is an FP7 project developing an ontology and database solutions for the data generated in the EU NanoSafety Cluster [2, 3]. This includes extracts of experimental data from, for example, cell line experiments, environmental toxicity studies, and high-throughput screening results. More important, however, is that this data is no longer static but can be queried and analysed. That is, to make the best use of this data, integration with other life science databases is needed, such as protein sequence database like Uniprot and compound databases such as ChEMBL [4] and PubChem [1]. Doing so allows us to test scientific hypotheses such as about the genotoxicity of metal oxides, whether chemically similar nanomaterials have similar bioactivities, or whether protein coronas contain preferably proteins involved in specific biological processes.

Semantic Web standards are an increasingly central interoperability layer linking experimental data to scientific knowledge. eNanoMapper has been working on extending the semantics of the database software to import and export data in a serialization based on the Resource Description Framework (RDF) and the eNanoMapper ontology. The RDF data is made available as dereferenceable data and via a SPARQL endpoint. These technologies are then used to support the research data management in the community. First, data completeness is checked by using SPARQL queries, thereby highlighting missing data. Second, the scientific questions predefined by the eNanoMapper project, such as mentioned earlier in this abstract, are supported by SPARQL queries aggregating the relevant data. Finally, the eNanoMapper RDF is enriched with links to other Linked Open Data Cloud resources (e.g. ChEMBL, PubChem) to support further nanosafety research.

References

1. Fu, G., Batchelor, C., Dumontier, M., Hastings, J., Willighagen, E., Bolton, E.: PubChemRDF: towards the semantic annotation of PubChem compound and substance databases. Journal of Cheminformatics 7(1), 34+ (Jul 2015)
2. Hastings, J., Jeliazkova, N., Owen, G., Tsiliki, G., Munteanu, C.R., Steinbeck, C., Willighagen, E.: eNanoMapper: harnessing ontologies to enable data integration for nanomaterial risk assessment. Journal of Biomedical Semantics 6(1), 10+ (Mar 2015)
3. Jeliazkova, N., Chomenidis, C., Doganis, P., Fadeel, B., Grafström, R., Hardy, B., Hastings, J., Hegi, M., Jeliazkov, V., Kochev, N., Kohonen, P., Munteanu, C.R., Sarimveis, H., Smeets, B., Sopasakis, P., Tsiliki, G., Vorgrimmler, D., Willighagen, E.: The eNanoMapper database for nanomaterial safety information. Beilstein Journal of Nanotechnology 6, 1609–1634 (Jul 2015)
4. Willighagen, E.L., Waagmeester, A., Spjuth, O., Ansell, P., Williams, A.J., Tkachenko, V., Hastings, J., Chen, B., Wild, D.J.: The ChEMBL database as linked open data. Journal of Cheminformatics 5(1), 23+ (May 2013)