Antreas Afantitis, Iseult Lynch

Digitising nanomaterials safety assessment - innovative nanoinformatics data workflows, models and tools for predictive nanomaterials (eco)toxicology

The need for *in silico* nanomaterials toxicity prediction

Nanotechnology is a key enabling technology, capable of delivering a wide range of technological breakthroughs across all EU priority sectors. Production of novel and emerging engineered and manufactured nanomaterials (NMs) is fundamental to advances in aeronautics, construction, electronics and consumer good. Although numerous benefits of NMs have been identified over the last decades, concerns are also arising as risk assessment lags behind product development, mainly because current approaches to assessing exposure, hazard and risk are expensive, time-consuming, and frequently involve testing in animal models. Additionally, the test guidance documents for generating regulatory-compliant data are still being updated for nanomaterials, whose surface reactivity and interactions with biomolecules and cellular machinery makes their assessment challenging compared to molecular chemicals. This leaves industry in the challenging position of having to provide data but how to generate the required data is not always clear.

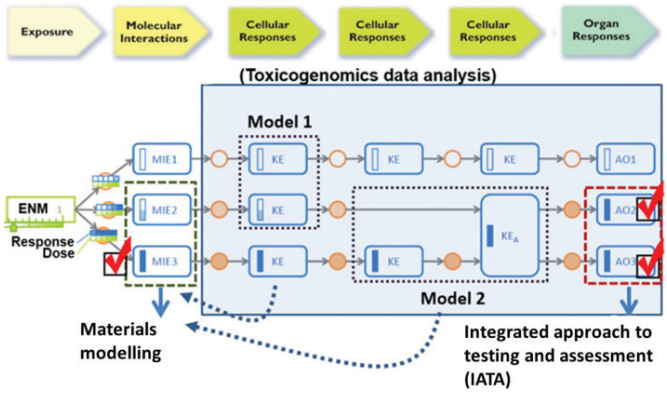


Fig. 1: Schematic representation of the NanoSolveIT Integrated Approach for Testing and Assessment (IATA) of NMs human and environmental risk.

Dr. Antreas Afantitis,
Novamechanics Ltd.
Coordinator of H2020 NanoSolveIT
afantitis@novamechanics.com

Prof. Iseult Lynch
School of Geography, Earth and Environmental Sciences,
University of Birmingham
Deputy Coordinator of H2020 NanoSolveIT
+44 121 414 5532
i.lynch@bham.ac.uk

DOI: .....

To overcome these challenges, the European Union H2020-funded project **NanoSolveIT** aspires to introduce a ground-breaking *in silico* Integrated Approach to Testing and Assessment (IATA) for the environmental health and safety of NMs, implemented through a decision support system packaged as both a stand-alone open software and via a Cloud platform (Figure 1). To achieve this, the NanoSolveIT consortium has an ambitious plan of model development, integration, testing and validation via case studies, and provision of a user-friendly graphical user interface and full training documentation to support implementation of the IATA and its uptake and application by industry and regulators.

The need for digitisation of nanosafety data

Development of *in silico* models is highly dependent on availability of large harmonised and high-quality datasets, and thus an important aspect of the NanoSolveIT project has been development of tools and workflows for pre-processing of experimental data, mining of existing literature and databases, and unique identifiers for NMs to ensure inter-operability of datasets. Where gaps in existing datasets are identified, innovative approaches to gap-filling have been developed, including generation of computational NMs descriptors directly from existing experimental data, or from compositional information encoded in the periodic table,[1] for example. Another approach underway is the generation of computational NMs with 1 nm resolution in size and shape, which can be utilised in "*in silico* experiments" to correlate properties with effects such cytotoxicity, genotoxicity or reproductive toxicity in the test organism *Daphnia magna*, for example.

A key challenge in integrating and harmonising nanosafety datasets from different publications and databases is ensuring that the data relate to the same NM. While chemoinformatics has developed efficient machine-readable ways of representing chemical structures for small molecules, such as the IUPAC International Chemical Identifier (InChI), the multi-component structures of NMs are more challenging to represent in a machine-readable format to enable linking of datasets for nanoinformatics and regulatory applications. To address this issue, NanoSolveIT partners in collaboration with the H2020 research infrastructure project NanoCommons (coordinated by Prof. Iseult Lynch) developed a proposal for how to extend the InChI to NMs,[2] and implemented an alpha version of a tool to generate NMs InChIs: http://enaloscloud.novamechanics.com/nanocommons/NInChI/. The first layer is a standard feature of all InChI-based notations indicating the version number. The

second layer describes subparts of the NM including core and shell materials but also ligands, impurities, dopants and specific linkers, as shown in Figure 2. For each of these components, the information given in this layer consists of the chemical composition, and additional layers to describe NM size, shape, crystal structure or chirality (where relevant) and surface features, defects, etc. The third layer then shows how these components are combined to build the final NM, starting from the inside and working outwards towards the reactive surface. Work is now underway, in collaboration with IUPAC and the InChI Trust to develop this into an official InChI standard for NMs.

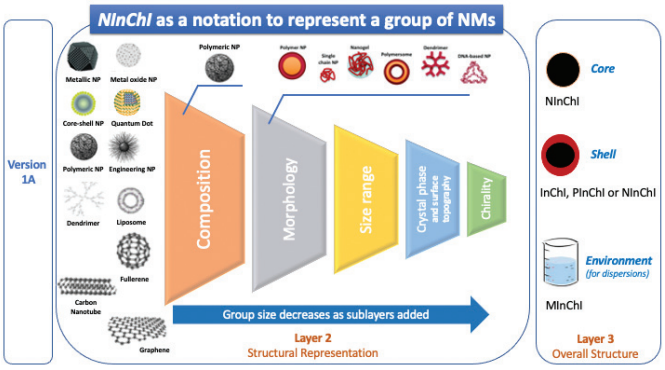


Fig. 2: Schematic illustration of the NInChI as a notation to represent a particular group of NMs, with the size of the group getting smaller as additional sub-layers and NM components (core, shell(s), surroundings in the case of NMs provided as dispersions). NInChI incorporates existing InChIs for small molecules, and extensions for polymeric coatings (PInChIs) and mixtures (MInChI) for particles in dispersion.

Another example of NanoSolveIT's innovative approaches to data integration and pre-processing is demonstrated via the creation of a unified collection of publicly available transcriptomics data on NM exposures. The original data collection step resulted in 140 data sets. Further selection was based on a manual quality assessment based on five criteria: 1. Three or more biological replicates included for statistical robustness; 2. Transcriptomics platform is a commercial gene expression microarray produced by Agilent, Affymetrix or Illumina; 3. Labelling of 2-color microarrays has considering dye swapping; 4. Non-normalized raw data is available; and 5. Supporting information reports all variables required for pre-processing. The data sets passing these criteria were pre-processed according to the standard NanoSolveIT protocol (described in detail in a series of 3 papers describing an optimised toxicogenomics workflow[3-5]), resulting in a data collection comprising 90 data sets relating to >50 unique NM versus control comparisons for differential expression, divided into four classes based on the number of timepoints and number of exposure concentrations. This dataset forms the basis of subsequent modelling work related to establishment of NMs Adverse Outcome Pathways (AOPs) and linking of these to exposure and toxicokinetics and toxicodynamics models such as physiologically based pharmacokinetic (PBPK) modelling.

Examples of NanoSolveIT *in silico* predictive nanosafety tools

The first deep learning model for NMs ecotoxicity was developed within NanoSolveIT, utilising chronic daphnia toxicity from paired studies of parental only exposure followed by recovery over 4 generations or continuous exposure, utilising both pristine and aged NMs, and salt-only versus natural organic matter containing media to mimic standardised versus environmentally realistic conditions. Toxicity was evaluated based on morphological features including changes in size, shape, abdomen, claw and heart defects, lipid deposits etc.[6] The predictive model is fully documented and made available to users via the dockerisation, via an Application Programming Interface, or via the user-friendly interface developed as part of the cloud platform, available at https://deepdaph.cloud.nanosolveit.eu/.

An example of gap-filling type model, that is less-reliant on experimental data through innovative extension of the data with computational descriptors, is the robust, validated and easily applicable model for the prediction of the cytotoxicity of metal oxide NMs developed in NanoSolveIT. The model was developed using a dataset containing 15 physicochemical and structural descriptors for the 24 metal oxide NMs included in the library, enriched with 62 atomic computational descriptors generated using the Isalos Analytics Platform and the Enalos+ nodes. Out of the 77 total descriptors used as input, 7 were deemed statistically significant, of which just two are experimental parameters (core and hydrodynamic size of the NMs), two are assay-related parameters (assay type (LDH or ATP) and exposure dose) and three are computational descriptors: the energy of the conduction band (E_C), the coordination number of surface metal atoms (Avg. C.N. Me atoms surface) and the force vector surface normal component of metal atoms (v⊥ Me atoms surface). E_C can be found from libraries of physicochemical descriptors (e.g., https://materialsproject.org/), and coordination number of surface atoms and the force vector surface can be calculated using molecular dynamics software (e.g., LAMMPS). The model allows read-across based on chemical similarity of specific Me_xO_y NMs and the use of the LDH or ATP assays to predict cytotoxicity. The curated dataset, including the computationally predicted descriptor values, is accessible and the model is publicly available as a webservice through NanoSolveIT Cloud Platform, https://cellviability.cloud.nanosolveit.eu/2/instructions.zul. The model is complemented with an application programming interface to make it available and easy to use programmatically, i.e., to implement into a workflow and to facilitate data exchange with databases to run the model.

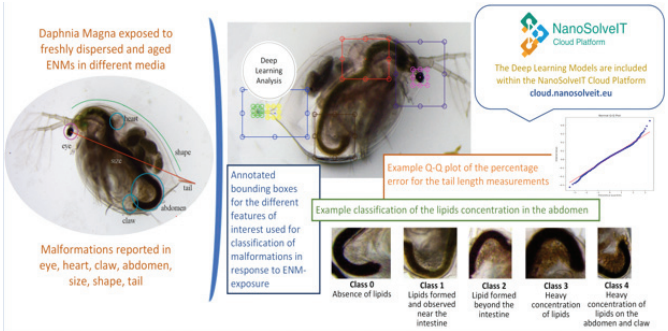


Fig. 3: NanoSolveIT's deep learning model for assessment of the chronic and multi-generational toxicity of NMs to *daphnia magna* which can distinguish pristine versus aged NMs. See https://doi.org/10.1002/smll.202001080 for further details, or access the model via the user-friendly interface: https://deepdaph.cloud.nanosolveit.eu/.

Links to the NanoSolveIT models and tools

All NanoSolveIT models are available via individual user-friendly interfaces from the project website: https://nanosolveit.eu/resources/tools-services/. At present each model is standalone, although the outputs of some of the models can already serve as the input datasets for other models. Our current work is building the integrated workflow, in collaboration with the research infrastructure project NanoCommons (https://infrastructure.nanocommons.eu/), which will initially integrate a human indoor air exposure model, a multi-compartment PBPK model extrapolated to humans, and a risk assessment and risk prediction tool. The final NanoSolveIT IATA will integrate a full suite of meta-scale models including computational NMs models, models to predict physicochemical descriptors of NMs and to gap-ill datasets, toxicogenomics and adverse outcome pathway models, Quantitative Structure Activity models, machine learning and artificial intelligence models and more.

References

[1] Papadiamantis *et al.*, Predicting Cytotoxicity of Metal Oxide Nanoparticles Using Isalos Analytics Platform. Nanomaterials 2020, **10(10)**, 2017; https://doi.org/10.3390/nano10102017

[2] Lynch *et al.*, Can an InChI for Nano Address the Need for a Simplified Representation of Complex Nanomaterials across Experimental and Nanoinformatics Studies? Nanomaterials 2020, **10(12)**, 2493; https://doi.org/10.3390/nano10122493;

[3] Kinaret *et al.*, Transcriptomics in Toxicogenomics, Part I: Experimental Design, Technologies, Publicly Available Data, and Regulatory Aspects. *Nanomaterials* 2020, **10(4)**, 750; https://doi.org/10.3390/nano10040750;

[4] Federico *et al.*, Transcriptomics in Toxicogenomics, Part II: Preprocessing and Differential Expression Analysis for High Quality Data. *Nanomaterials* 2020, **10(5)**, 903; https://doi.org/10.3390/nano10050903;

[5] Serra *et al.*, Transcriptomics in Toxicogenomics, Part III: Data Modelling for Risk Assessment. *Nanomaterials* 2020, **10(4)**, 708; https://doi.org/10.3390/nano10040708;

[6] Karatzas *et al.*, Development of Deep Learning Models for Predicting the Effects of Exposure to Engineered Nanomaterials on Daphnia magna. Small, 2020, **16 (36)**, 2001080; https://doi.org/10.1002/smll.202001080.

Antreas Afantitis, PhD, MBA
CEO NovaMechanics Ltd.
Coordinator H2020 NanoSolveIT project

Antreas received his B.Sc in Chemical Engineering (2003), M.Sc in Computational Mechanics (2004) and PhD in Computational & Organic Chemistry (2006) from the School of Chemical Engineering at the National Technical University, Athens (NTUA) in Greece.

He is the founder and director of NovaMechanics Ltd, a Cypriot informatics company, and is responsible for the overall management, strategic direction, growth and financial control. He has a strong scientific background in the field of chemoinformatics, bioinformatics, nanoinformatics, modelling, simulation and medicinal & materials chemistry. His scientific work has been published in over 80 original research articles and reviews in international peer reviewed journals. He has been included in the most productive authors in the field of Chemoinformatics (bibliometric analysis performed by Prof. Peter Willet) and according to Google Scholar his h-index is 29. He has successfully led the development and the implementation of state of the art information technology systems (databases, web services, custom made scientific software development) in NovaMechanics Ltd for solving Cheminformatics, Bioinformatics, Nanoinformatics, Modelling, Simulation and Big Data analysis problems in drug discovery, nanosafety and more.

Prof. Iseult Lynch
Professor of Environmental Nanosciences
University of Birmingham
Deputy Coordinator H2020 NanoSolveIT project

Iseult has a Bsc. and PhD in Chemistry, and an MBA, from University College Dublin, Ireland and is Chair Professor in Environmental Nanosciences at the University of Birmingham (UoB), where she leads an international team of researchers assessing the environmental chemistry and toxicity of emerging pollutants. She has a very broad overview of all aspects of chemicals safety assessment, with a research focus on the extension of REACH and other regulatory frameworks to fully account for nanomaterials surface reactivity and NMs interactions with biomolecules (proteins, metabolites, natural organic matter, pollutants etc.), and the implications for both fate and effects (toxicology).

She is a Fellow of the Royal Society of Chemistry, and an Associate Editor for the RSC journal Environmental Science: Nano. She is an advisor to NovaMechanics and works closely with the NovaMechanics management team in developing and delivering strategic projects and identifying key opportunities for regulatory utilisation of their tools and approaches.

NanoSolveIT - a H2020 nanoinformatics project addressing nanomaterials safety

Innovative Nanoinformatics models and tools: towards a Solid, verified and Integrated Approach to Predictive (eco)Toxicology (NanoSolveIT) is a H2020 project running from January 2019 to February 2023. NanoSolveIT's ambition is to advance nanoinformatics well beyond the state-of-the art by developing and implementing innovative modelling techniques and tools that will be integrated within the NanoSolveIT *Integrated Approach for Testing and Assessment* (IATA) and packaged as a sustainable interoperable product, the NanoSolveIT e-platform. Our vision is that the NanoSolveIT IATA will become an essential element for supporting industrial and regulatory risk assessment and governance of nanomaterials, including identification of individual *nanoforms* and *sets of nanoforms* under REACH's 2019 updated guidance.

NanoSolveIT is funded through the European Union Horizon 2020 Programme (H2020) under grant agreement nº 814572. https://nanosolveit.eu/ Follow @NanoSolveIT

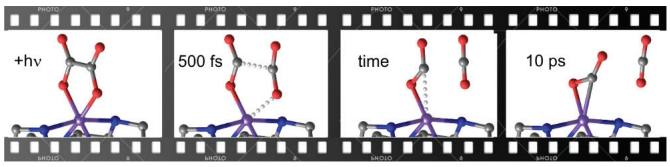
Hot paper

Wettbewerb: Wo schwingt das Wasser in der Organik?

Wenn ein einzelnes Wassermolekül an ein organisches Molekül andockt, verändern sich seine Streckschwingungsfrequenzen in charakteristischer Weise - je nach Stärke und Art der Wasserstoffbrücke. Im Rahmen des **DFG-Graduiertenkollegs BENCh (DFG GK 2455) der Georg-August Universität Göttingen** soll ab diesem Jahr ein **Mikrohydratations-Wettbewerb unter Theoriegruppen weltweit** veranstaltet werden, um dies verlässlich vorherzusagen - sei es mit skalierten harmonischen Vohersagen, mit unterschiedlichsten anharmonischen Modellen, oder gar mit maschinellem Lernen. Neben einem beträchtlichen Trainingsdatensatz aus der experimentellen Literatur sollen im Rahmen des GKs im Sinne eines echten Blindwettbewerbs neue experimentelle Testsysteme beigesteuert werden - auch noch während die Theoretiker*innen ihre Vorhersagen machen. Der Start wurde mit einer ungewöhnlich systematischen anharmonischen Resonanz bei Keton-Monohydraten gerade vorbereitet (https://doi.org/10.1021/acs.jpclett.0c03197, siehe SI). Weitere Schritte und **Details** werden im Laufe des Jahres unter **bench.uni-goettingen.de** angekündigt.

Kontakt: **Martin Suhm** (msuhm@gwdg.de)

Hot paper



Carbon Dioxide Binding – Insights from Ultrafast Spectroscopy

The transformation of CO_2 to value-added products and fuels will become a strategically important contribution from the Chemical Sector to meet the goals expressed in the Paris Agreement and the European Green Deal. Sustainable CO_2-conversions may be catalyzed by earth-abundant transition metals (TMs) like iron. Understanding the molecular and electronic structures that govern the binding of CO_2 to TMs is thus of critical importance. The photochemistry of oxalato-TM complexes can serve as an exquisite model for studying these aspects in detail [1]. **Optical excitation cleaves the oxalate moiety heterolytically into a CO_2 molecule**, that leaves the ligand sphere, **and a carbonite dianion**, CO_2^{2}, that is retained by the metal as a "redox active" ligand. Accordingly, the photolysis of a ferric precursor must be expressed with three limiting electronic structures of the photoproduct: $Fe^{3+}–C_2O_4^{2-} + h\nu \rightarrow [Fe^{3+}–CO_2^{2-} \leftrightarrow Fe^{2+}–CO_2^{\bullet -} \leftrightarrow Fe^{+}–CO_2^{0}] + CO_2$. Very recently [2] **ultrafast mid-IR spectroscopy in combination with DFT** showed that these three forms give rise to distinct TM–CO_2 binding modes: (i) a "side on"-mode with a carbonite ligand, (ii) a bent "O-end-on"-mode with a radical anion ligand, and (iii) a linear "O-end-on"-mode with a neutral ligand. By tuning the ligand sphere the ground-state spin of the Fe–CO_2 complex can be selected, which in turn, allows for steering the binding mode in a desired direction. **Can the spin thus serve as a toggle for O vs C centered primary reactivity of a CO_2 attached to a metal?** A provocative but intriguing question, which will be addressed in the future with time-resolved spectroscopies. Research on such light controlled processes of TM complexes is the subject of the **DFG-funded Priority Program, SPP 2102** [3].

[1] Steffen Straub, Paul Brünker, Jörg Lindner, Peter Vöhringer: An Iron Complex with a Bent, O-Coordinated CO_2 Ligand Discovered by Femtosecond Mid-Infrared Spectroscopy, *Angew. Chem. Int. Ed.* 2018 **57**, 5000-5005 (http://dx.doi.org/10.1002/anie.201800672).

[2] Steffen Straub, Paul Brünker, Jörg Lindner, Peter Vöhringer: Spin-Controlled Binding of Carbon Dioxide by an Iron Center: Insights from Ultrafast Mid-Infrared Spectroscopy, *Angew. Chem. Int. Ed.* 2021 **60**, accepted Oct-2020, (http://dx.doi.org/10.1002/anie.202012739).

[3] https://www.spp2102.uni-mainz.de

Contact: **Peter Vöhringer** (pvoehr@uni-bonn.de)