

Making project data available through eNanoMapper database: NANoREG, Nanoreg2,caLIBRAte and Gracious

Nina Jeliazkova, Ideaconsult Ltd., Sofia, Bulgaria

Presented by Georgia Tsiliki, ATHENA Research Center



eNanoMapper overview



ENM Summary & data solutions:

- FP7 eNanoMapper A Database and Ontology Framework for Nanomaterials Design and Safety Assessment
- Grant Agreement: 604134
- Duration: 1 Feb 2014 31 Jan 2017; 8 partners



- Open source database and web application
 - Builds upon a Chemical structure database with support for substances
 - The data model supporting experimental data is capable of representing all endpoints of regulatory interests and other types of data.
- eNanoMapper ontology; developed by an experienced team at EBI. Existing ontologies are reused;
- Tools to process and import data. Export in various formats
- Searchable; Free text search based on ontology
- eNanomapper modelling; Integration of data analysis tools via API led by NTUA. WEKA, R python routines uploaded to Jaqpot platform
- Flexible data hosting architecture



Organising the nanosafety data

Challenges

- Diverse data sources
- Diverse data input formats
- Different data organization
- Diverse modelling tools
- Approach:
 - Enable mappings!
 - i.e. eNanoMapper

Physico-chemical identity

Different analytic techniques, manufacturing conditions, batch effects, mixtures, impurities, size distribution, differences in the amount of surface modification, etc.

Biological identity

Wide variety of measurements, toxicity pathways, effects of ENM coronas, modes-of-action, interactions (cell lines, assays).

Processes requiring information

From raw data (science) to study summaries for regulatory purposes; linking with experimental protocols; risk assessment; grouping, safety-by-design

Support for data analysis

Requires "spreadsheet" or matrix view of data. The experimental data in the public datasets is usually not in a form appropriate for modelling (merging multiple values, conditions, similar experiments into matrix form is a challenge).



Public eNanoMapper database

https://data.enanomapper.net

ME > ARTICLES > VOLUMES > VOLUME 6 > 10.3762/BJNANO.6.165

The eNanoMapper database for nanomaterial safety information

Nina Jeliazkova¹ ∞, Charalampos Chomenidis², Philip Doganis², Bengt Fadeel³, Roland Grafström³, Barry Hardy⁴, Janna Hastings⁵, Markus Hegi⁴, Vedrin Jeliazkov¹, Nikolay Kochev^{1,6}, Pekka Kohonen³, Cristian R. Munteanu^{7,8}, Haralambos Sarimyeis², Bart Smeets⁷, Pantelis Sopasakis^{2,9}, Georgia Tsiliki², David Vorgrimmler¹⁰ and Egon Willighagen⁷

- Mostly literature data + partial content provided by MODENA and MARINA projects; links to external DB
- Free text search <u>https://search.data.enanomapper.net/enm</u>

G	Secure http	ps://data.enanomapper.net/bundle				Q # 0 0
owing	9 bundles (1 to 9)				First P	wwitten 2 News, Lard
•	Bundle ID 0	Name 0	Ver ¢	Description 0	Source 0	License & Maintainer 🍳
	00000000-0000- 0000-0000- 000000000000	NanoWiki	<u>v2</u>	Nanomaterials, physicochemical characterisations and toxicity data, imported via NanoWiki RDF dump	NanoWiki http://dx.doi.org/10.6084/m9.figshare.1330208 cf	License t http://orcid.org/0000- 0001-7542-0286 cf
	00000000-0000- 0000-0000- 00000000002	OECD Harmonized Templates import test	<u>×1</u>	Demonstrate import of a nanomaterial and endpoints data form IUCLIDS *.iSz file (OECD HT)	Multi-Walled Carbon Nanotubes (MWCNT), synthetic graphite in tubular shape http://www.shape.sha.eu/ceasitered/data/dossiers/D155-b281d1a0-c6d8-5dcf- #044-00144f67d031/Ad0R-cd53254s-7b90-4a11-042d 72005c821048_D1555-b2014a0-c6d8-354*e044-00144f67d031.html cf	License : www.ideaconsult.net
	00000000-0000- 0000-0000- 000000000003	Protein Corona Fingerprinting Predicts the Cellular Interaction of Gold and Silver Nanoparticles	¥7	Demonstrates import, display and search of coated Ag and Au nanoparticles with large number of physicochemical characterisations data and biological responses. Serves as a test case for NanoQSAR (eNanoMapper WP4) modelling activities.	10.1021/nn406018q http://ode.acs.org/dol/abs/10.1021/nn406018g cf	License :
	<u>91cad054-4de8-4dc9- a8e4-20564e7eada7</u>	In Vitro data from FP7 MARINA project (KI)	<u>×1</u>	Publicly released data from publication 10.1371/journal.pone.0127174	Comprehensive In Vitro Toxicity Testing of a Panel of Representative Oxide Nanomaterials: First Steps towards an Intelligent Testing Strategy http://dx.doi.org/10.1371/journal.gon.0127174 cf	Licenze : P32 - KI
	<u>13e147c3-ccb5-11e5-</u> <u>246b-00ee7350bfa7</u>	Sigma-Aldrich Products	¥1	List of nanomaterials that you can buy with Sigma- Aldrich.	NanoWiki http://dx.doi.org/10.6084/m9.figshare.1330208 cf	License n http://orcid.org/0000- 0001-7542-0286 c*
	646ab3a4-ccb5-11e5- 946b-80ee7350b/a7	JRC Representative Nanomaterials	<u>×1</u>	Repository of representative nanomaterials with the purpose that many laboratories can measure physicochemical and biological properties on them.	NanoWiki http://dx.doi.org/10.6084/m9.figthare.1330208 cf	License : http://orcid.org/0000- 0001-7542-0286 :*
	804608b3-ccb5-11e5- 946b-80ee7350bfa7	Crystallography Open Database	¥1	List of nanomaterials in the Crystallography Open Database.	NanoWiki http://dx.doi.org/10.6084/m9.figshare.1330208 cf	License : http://orcid.org/0000- 0001-7542-0286 :*
	ee0cdb4c-ccel-11e5- 946b-80ee7350bfa7	ArrayExpress	¥1	ArrayExpress	NanoWiki http://dx.doi.org/10.6084/m9.flashare.1330208 ో	License : http://orcid.org/0000- 0001-7542-0286 5*
	62b0a66b-cd01-11e5- a67e-80ee7350bfa7	MODENA	<u>¥1</u>	MODENA TD1204 COST ACTION dataset	MODENA TD1204 COST ACTION	License :

eNanoMapper database instances



search.data.enanomapper.net

B eNanoMapper	×	100		Cardinal Control of Co	and the second second		
← → C (@ http:	s://apps.id	eaconsult.net/na	anoreg1				★]0 0 i
	lome	Search •	Data collections •	Data upload *	For developers •	Help •	Login

- Public
 - FP7 NANoREG (2013-2017)
- Restricted
 - FP7 MARINA (2011-2015)
 - NanoGenotox (2010- 2013)
 - FP7 NanoTest (2008-2012)
 - ENPRA
 - (import ongoing)

	NAN SREG	NANoREG - eNanoMapper database	
		Release 01 02 2018 (content as of TH/O DB 7.11 2017)	
	frei	E beatlange bashada an an an	(<u>db</u>)(<u>e</u>)
	TIO C Secure http	s://appsideaconsult.net/marina/ui	* 0
	sea Home	Search Data collections Data upload For developers Help	[h1070501
	C electron	FP7 MARINA : Managing Risks of Nanom	aterials
	← → C ■ Sec	/re https://aps.ideaconsult.net/hanogenotos/ui	\$
	-	Home Search * Data collections * Data upload * For developers * Help *	
oMapper x		Release Of 102.001 Mp. Insergendut as	
C @ https://apps.ic	eaconsult.net/nanotest/ui		\$ 0 0 i
Home	Search • Data collections •	Data upload • For developers • Help •	Logun
	NaneTEST	NANOTEST	
	JRCNM02003a	Search.	
	Search by identifier by sthytichem s	animeters.or.biolopic.at.effects bx.composition by.citation Browse Uplead	
	elitanolitapper database The elitanolitapper printippe database in y within the <u>EULIPEC attaination</u> Provides support for upload, search and re The elitanolitapper prototype database in y	er d the computational influenticulars for toxicultypical bits management of engineered merintensis, shoutiged 12. 13. 14. Second and informationals and experimental data through a NEST with second and and a second s	-

Bublemann doi 10 1000/amann 6 165, mf

NANoREG – eNanoMapper database



search.data.enanomapper.net/nanoreg

0

Home eNanoMapper NANoREG NanoReg2 caLIBRAte GRACIOUS

Nano safety data

NANoREG Results Repository Experimental data

In NANOREG over 85 institutional partners from EU member states, associated states, the Republic of Korea and Brazil collaborated in developing reliable, reproducible and relevant methods for testing and assessing the effects of nanomaterials on human health and environment in a regulatory context.

The duration of the project was 48 months; the end date was 28 February 2017. The budget was approximately fifty million euro. Ten million was provided by the EU (FP7 programme) and forty million by member states, regions, partners and other sources.

As -probably the first - nanosafety project, all the information and knowledge generated under the umbrella of the NANOREG project has been made available for other parties directly after the ending of the project. The NANOREG Results Repository serves as a central point of access to all the results of the project by providing links to all relevant documents, datasets and other information of interest.

A large part of the NANOREG experimental nano-EHS data is available via the **NANOREG** eNanoMapper database. It contains data gathered via TNO Data Entry Tool (snapshot as of 19.05.2017). The coming months more data will be added. An **user guide** is available.

For NANOREG partners It was mandatory to carry out the experiments with NANOREG core materials. For all these core materials a complete state-of-the-art physicochemical characterization was done considering key end-points in the OECD WPMN sponsorship program.

Dispersion SOPs and minimum requirements for characterization were established and laid down in a Guidance Document thus supporting reliability and comparability of data. Experiments had to be performed in accordance with the Guidance Document. During the course of the project, those fundamental requirements were further refined.

The data in the data base can be further analyzed by various tools.

License

If not indicated otherwise, the information in the NANoREG Results Repository is available under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. Information on commercial use and attribution can be found here.

NANoREG Results repository read more NANoREG

Experimental data

zinc oxide

NANoREG - eNanoMapper database

Search



Data transfer agreement NANoREGeNanoMapper (2016)

Publicly accessible since March 2017

Open data license

Content imported from

- TNO SQL database
- Excel files

NanoSafety data : top level view



https://search.enanomapper.net

*	Home	NanoMapper	NANoREG	NanoReg2	caLIBRAte	GRACIOUS	
eNanoMapper database							
The eNanoMapper database is part of the computation within the EU FP7 eNanoMapper project. More information	al infrastructure fo ation:	r toxicological da	ta manageme	nt of engineer	ed nanomateri	als, developed	
Data content				Data acc	ess		
eNanoMapper							
NANoREG				Public			
FP7 NanoTest				eNanoMa NANoREC	pper }		
FP7 NanoGenotox							
FP7 MARINA							
H2020 NanoReg2							
H2020 caLIBRAte							
H2020 GRACIOUS							
Acknowledgements							
 This project has received funding from the Europea development and demonstration under grant agree 	n Union's FP7 for re ment no 604134 eN	search, technolo anoMapper	gical				
FP7 NANoREG collaboration							
H2020 caLIBRAte collaboration							
 This project has received funding from the Europea innovation programme under grant agreement 6462 	n Union's Horizon 2 221 NanoReg 2	020 research and	1				

Search data integration



https://search.enanomapper.net



How to retrieve the data ?



Free text and faceted search

C 🛈 https://apps.ideaconsult.net/nanodata,	/nanoreg2/?	&ui=%257B%2522owner_name%2522%253A%257B%2522values%2522%253A%255B%255D%257D%252C%2522r	eference_owner%2522%253A%257B%2	522values% 😭 🛈 🖡
NanoReg ² Home Sear	ch Su	mmary Data collections Data templates Help [n10705013]		
NanoReg2 database		Search	l .	
License information: FP7 NANOREG FP7 MAR	INA FP7 I	NanoTest FP7 NanoGenotox About		
Projects (23267)	Hite lie	t Selection		Expor
ilter_	A549	titanium oxide Genetic toxicity in vitro		Clear
ANOREG NanoGenotox NanoTest AND	University			(Enstrand
Study providers (23267)	< 1 > dis	PROMOTORO (NM-100 (TiO2 E0.1E0 pm)) titenium exide percenticle		
	nişer	CORE (1):		L
Nanomaterial type (92816)		Results:P-CHEM.Crystalline phase, P-CHEM.Surface chemistry, P-CHEM.Particle size distribution CHEM.Batch Dispersion guality, P-CHEM.Specific surface area, TOX.Cell Viability, TOX.Immunot	n (Granulometry), P-CHEM.Aspect oxicity, TOX.Genetic toxicity in vitr	Selecte
Nanomaterial (22829)		toxicity - inhalation more		C11
Protocols (67978)		Material Composition Studies		filter Section
Method (55885)	-	IPCNM01001a (NM-101 (TiO2 6 nm)) titanium ovide nanonarticle		e
TOX (59034)	orşm	CORE (1):		
P-CHEM (8818)		Results:P-CHEM.Surface chemistry, P-CHEM.Crystalline phase, P-CHEM.Particle size distribution Dispersion quality, P-CHEM.Specific surface area, TOX.Immunotoxicity, TOX.Genetic toxicity in	n (Granulometry), P-CHEM.Dustine vitro, TOX.Cell Viability, TOX.Barri	ess, P-CHEM.Batch er integrity,
ECOTOX (462)		TOX.Oxidative Stress, ECOTOX.Short-term toxicity to aquatic inverterbrates, TOX.Repeated dos CHEM.Aerosol characterisation	e toxicity - inhalation, P-CHEM.Ze	ta potential, P-
Cell (55294)		more		
Species (8846)		material composition Studies		Add to Selection
Execute rate (210)	aufer	JRCNM01003a (NM-103 (TiO2 24.7 nm)) titanium oxide nanoparticle CORE (1):		e
Exposure route (510)		Results:P-CHEM.Surface chemistry, P-CHEM.Crystalline phase, P-CHEM.Particle size distributio	List of data	
Facets or filters		more		
41 4		Material Composition Studies	resources w	ith direct
that permit easy	100	JRCNM01004a (NM-104 (Titanium Dioxide)) titanium oxide nanopartic	links to DR	
refinement of	-	CORE (1): <u>Titanium dioxide</u> Results: P-CHEM.Specific surface area. TOX Repeated dose toxicity - oral. P-CHEM.Crystallite a		
Tennement Of		(Granulometry), TOX.Genetic toxicity in vitro, P-CHEM.PC_THERMAL_STABILITY_SECTION, TOX	C.Cell Viability, P-CHEM.IMPURITY_	SECTION
search		Material Composition Studies		Add to Selection
		JRCNM01003a (NM-103 (Titanium Dioxide)) titanium oxide nanoparticle	e	e
		Results:TOX.Repeated dose toxicity - oral, TOX.Cell Viability, TOX.Genetic toxicity in vitro. P-C	HEM.Particle size distribution (Gra	nulometry), P-
		CHEM.Crystallite and grain phase, P-CHEM.Specific surface area, P-CHEM.PC_THERMAL_STABIL more	ITY_SECTION, P-CHEM.IMPURITY	_SECTION
		Material Composition Studies		Add to Selection

How to retrieve and download the data ?



Export example

GRACIOUS ×			×
\leftrightarrow \rightarrow C $\textcircled{0}$ https://apps.ide	aconsult.net/nanod	ata/gracious/index.html?&ui=%2578%2522owner_name%2522%253A%2578%2522values%2522%253A%2558%255D%257D%252C%2522reference_owner%2522%253A%2578%252 🖈 💿	
	Home D	ata collections ▼ Summary Data templates ▼ Help ▼ [n10705013]	
gracious	GRACIOUS	5 database Search	
	H2020 GRACIC	US project About	
Projects (1369)		Hits list Selection Exp	ort
• Study providers (1369)		
Nanomaterial type (13	369)	Filtered entries Selected entries	
• Nanomaterial (66392)		Select export type	
Protocols (4085)		Materials and all study results Study results Material identifiers Material composition	
• Method (1088)			
• TOX (4011)	د	JSON CSV TXT XML XLSX RDF 80040 ISA	
Filter_		Download filtered entries as TSV	
Cell Viability	1		
Oxidative Stress	1		

- To download the search results specify "Filtered entries" and "Study results".
- Click on the TXT icon. The download button caption will change to
 - "Download filtered entries as TSV" (tab separated values)
- Note that only limited set of fields are exported in the TXT format.
- The JSON and XML formats contain the full set of fields.

Mapping internal database functions to the external world



- Application Programming Interface (API): a way computer programs talk to one another. Can be understood in terms of how a programmer sends instructions between programs.
- Access the database via any programming language, Workflow systems, Data analysis tools (*R*, JavaScript, Java, Ruby used by eNanomapper partners)
- eNanoMapper Tutorials:

http://www.enanomapper.net/enm-tutorials https://github.com/enanomapper/tutorials

How to retrieve and analyse the data?





Data entry issues

- Input data files: >1,000 Excel files roughly following
 - IOM templates (most projects except NANoREG)
 - NANoREG / JRC *ISA-TAB-logic* files
 - NANoREG / JRC are NOT ISA-TAB-NANO compliant
- Assays naming : multiple names per assay
 - Colony Forming
 - Colony Forming Assay
 - Colony Forming Efficiency Assay
 - CFE
 - Cell viability Colony Forming Efficiency Assay
- Naming materials, cells , methods, etc



Mapping terms: ontology

An ontology is a controlled vocabulary enhanced with relationships between terms.

<u>Usage</u>

- Harmonize data from several sources
- Analysis (logical inference, semantic search)

Examples:

- Is Cytotoxicity an endpoint, or EC50 is also an endpoint?
- Is TEM a protocol, or a technology?
- "Water Fish D. Rerio" assay – is it the same one as "Zebrafish Embryo Toxicity Test" ?



Hastings et al. Journal of Biomedical Semantics (2015) 6:10 DOI 10.1186/s13326-015-0005-5



RESEARCH

Open Access

eNanoMapper: harnessing ontologies to enable data integration for nanomaterial risk assessment

Janna Hastings¹, Nina Jeliazkova², Gareth Owen¹, Georgia Tsiliki³, Cristian R Munteanu⁴⁵, Christoph Steinbeck¹ and Egon Willighagen⁵

eNanoMapper overview



Mapping data organisations



http://ambit.sourceforge.net/enanomapper/templates/convertor_how.html

eNanoMapper overview



Mapping spreadsheet content into the data model



What is in a eNanoMapper database instance:



Composition

		Home :	Search V Data	collections •	Data uplo	ad • For developers	• Help •			
	 Search na 	nomaterials by identifiers					Name Identifier	Reference	NM type	
		1	1							
	Nanomateria	Advanced search	Download	_	_			_	_	
	Showing from	m 1 to 10 in pages of 10	🔹 substances 📹 [Previous Next	•			Filter		
		Substance Name	Substance UUID 🍦	Substan	се Туре	Public name	Reference substance \$ UUID	Owne	er ÷	Info
	□ · 1 · ►	Multi-Walled Carbon Nanotubes (MWCNT), synthetic graphite in tubular shape	<u>IUC5-5f313d1f-</u> ®	mono constit substance	tuent	Multi-Walled Carbon Nanotubes (MWCNT), synthetic graphite in tubular shape		<u>Ideaconsult Lto</u> <u>Bulgaria</u>	<u>d, / Sofia /</u>	
	- 2 - T	<u>G15.AC</u>	e <u>ECSV-bc77c03d-</u>	NPO_401		G15.AC	<u>FCSV-50cca421-</u> ම	Protein Corona Fingerprinting the Cellular Int Gold and Silver Nanoparticles.co	Predicts teraction of C CSV	Classification = Anioni
	- 3 - 9	<u>G15.AHT</u>	<u>FCSV-8f5cd32a-</u> ල	NPO_401		G15.AHT	FCSV-50cca421 €	Protein Corona Fingerprinting the Cellular Int Gold and Silver Nanoparticles.c	Predicts teraction of C CSV	Classification = Cation
ating		Composition Composition Purity of IUC Subsi	name: UUID: FCSV-8f5c tance:	d32a-3350-300	0b-91d0-87000	ee5d7ee		1		
		Туре	Name 🎄	EC No. 👙	CAS No.	Typical concentration	Concentration	on ranges	÷	Structure
		Coating 6-Am	ino-1-Hexanethiol			0 % (w/w)	0 % (w/w)	0 % (m/m)	<u>Also contair</u> <u>in</u>	
		Core Au			7440-57-5	0 % (w/w)	0 % (w/w)	0 % (w/w)	<u>Also contain</u> <u>in</u>	ned

What is in a eNanoMapper database instance:



Tox/Ecotox/Env fate

🚱 eNanoMapper	×			· · · · · · · · · · · · · · · · · · ·		-					Nra 🗖 🗖 🗙
$\leftrightarrow \rightarrow \mathbf{C}$ \bullet Secure	https://apps.ideacons	sult.net/nanoreg	1/substance/NNRG-a51b2e	58-4105-9643-3016-3	3f4b431171e2/study						\star 🛛 🖬 :
OENN	1 Home	Search •	Data collections •	Data upload •	For developers *	Help 🔻					<u>Log in</u>
▲ > Search nanom	naterials by identifiers	NNRG-a51t	<u>52e58-4105-9643-3016-3</u>	8f4b431171e2 \ Stu	l <u>dy</u> ⊚ N	Name	al identifier © E	xperiment reference	e	Search	
IUC Substance	Composition	Tox (130)	P-Chem (157)	Eco Tox (4)						Expar	ıd all Collapse all
JRCNM01001a											
NPO_1339 Immun	otoxicity (22)										*
7.5.2 Repeated do	se toxicity - inhalati	ion (12)									•
7.6.1 Genetic toxic	ity in vitro (30)										•
ENM_0000044 Bar	rier integrity (3)										•
ENM_0000037 Oxi	dative Stress (7)										•
ENM_0000068 Cell	Viability (56)										•



All projects data statistics



Physchem characterisation



All projects data statistics

Bioassays





Future plans









Future plans







www.patrols-h2020.eu

Future plans



Type of NPs:

- Engineered NPs (ZrO₂, CeO, Sb-Sn Ox, SnO)
- Process-generated NPs (atmospheric plasma spraying)

Type of data:

- Physical-chemical characterisation: particle size, distribution, solubility, zeta potential...
- Toxicological characterisation: ALI, submerged-cell, simulated body fluid, ROS generation.



Engineered NPs



Process-generated NPs









Thank you! Questions?

NanoReg2

