

# eNanoMapper database, search tools and templates

Nina Jeliaskova, Nikolay Kochev



IdeaConsult Ltd.

Sofia, Bulgaria

[www.ideaconsult.net](http://www.ideaconsult.net)



- eNanoMapper database: data model, technology;  
NANoREG data transfer examples
- Search tools: free text, chemistry, semantic; API access
- I/O support: ISA & Excel templates

## CONTENT

# eNanoMapper summary

**FP7 eNanoMapper - A Database and Ontology Framework for Nanomaterials Design and Safety Assessment**

**Grant Agreement: 604134,  
1 Feb 2014 – 31 Jan 2017**

**Solutions available**

**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.**

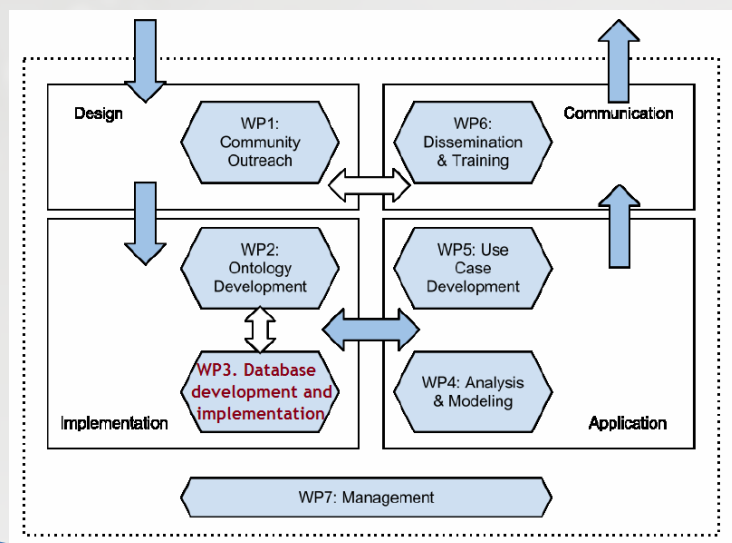
**eNanoMapper ontology; developed by an experienced team at EBI. Existing ontologies are reused;**

**Tools to process and import data. Export in various formats**

**Searchable; Free text search based on ontology**

**Integration of data analysis tools via API**

**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**

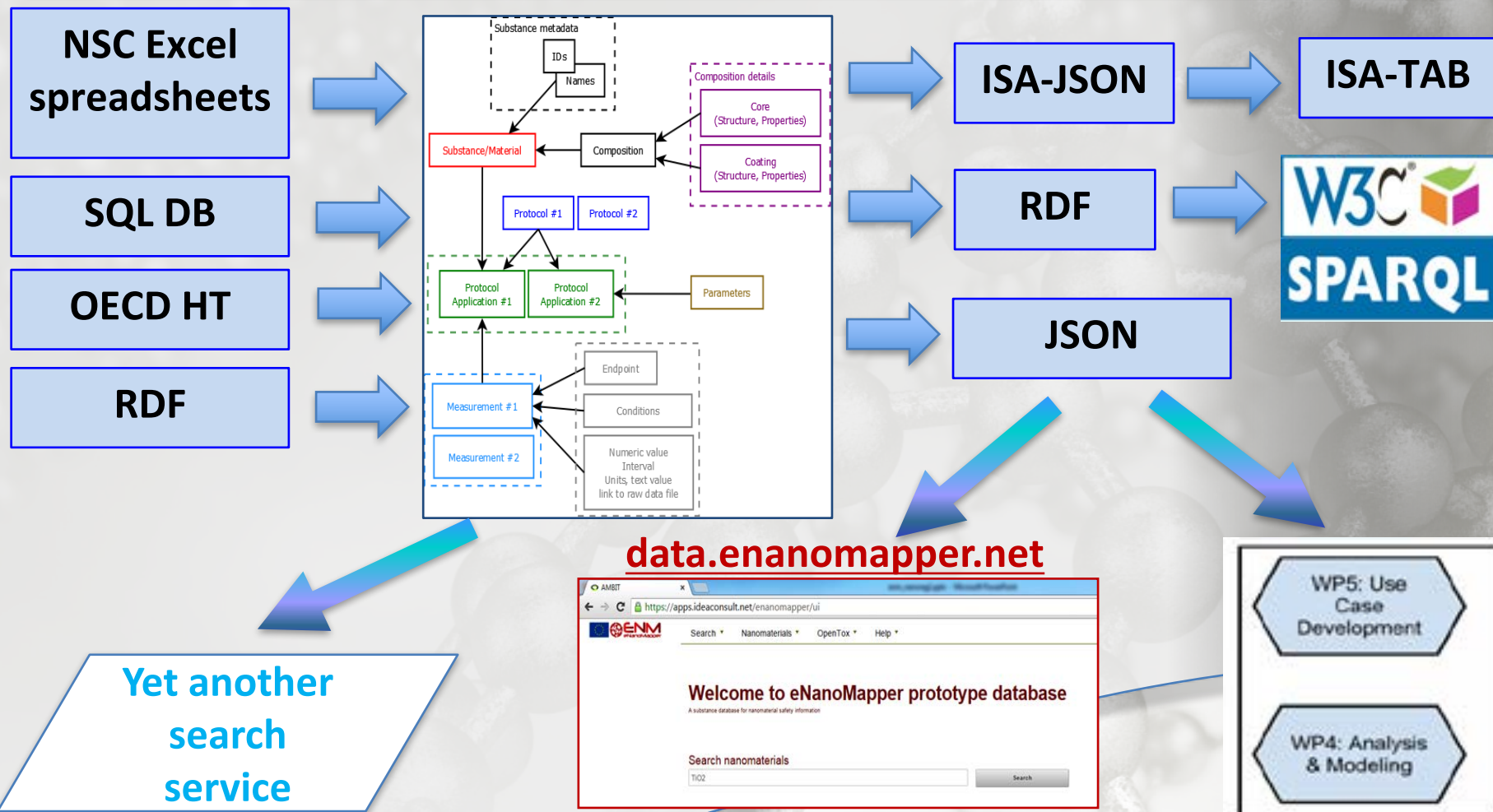
From raw data (science) to study summaries for regulatory purposes; linking with experimental protocols; risk assessment; grouping, safety-by-design

- **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).

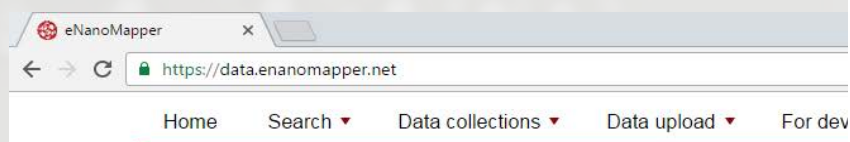


# The common eNanoMapper data model : enables conversions



# data.enanomapper.net

N. Jeliaskova, et al. “The eNanoMapper database for nanomaterial safety information,” *Beilstein J. Nanotechnol.*, vol. 6, pp. 1609–1634, Jul. 2015.



## eNanoMapper prototype database

A substance database for nanomaterial safety information

### free text search

Search by [identifier](#) | [physchem parameters or biological effects](#) | [by composition](#) | [by citation](#) | [Browse](#) | [Upload](#)

#### eNanoMapper database

The eNanoMapper prototype database is part of the computational infrastructure for toxicological data management of engineered nanomaterials, developed within the [EU FP7 eNanoMapper](#) project. Provides support for upload, search and retrieval of nanomaterials and experimental data through a REST web services API and a web browser interface.

The eNanoMapper prototype database is an open source web application, which can be [downloaded](#), installed and hosted by individual researchers or labs, and as such presents an open distributed platform for nanomaterials data management.

Publication: [doi:10.3762/bjnano.6.165](#)

eNanoMapper FP7 #604134. This project has received



# Implementation

- The database structure has two major concepts:
  - Substances, substance compositions, chemical structures
  - Experimental results (P-CHEM, ECOTOX, TOX, ENV-FATE)
- A generic description of any measurement. Does not specify what to record to describe particular experiment.
  - This information comes from NANoREG templates, IUCLID5 files, etc.
- The database software is based on an open source project <http://ambit.sf.net>
  - developed by eNanoMapper partner Ideaconsult since 2005, most recently : CEFIC LRI AMBIT tool for read across.
- The data model is capable of representing all endpoints of regulatory interests and other types of data.





# NANoREG data transfer (ongoing)

ENM eNanoMapper

Home Search Data collections Data upload For developers Help

Search nanomaterials by identifiers ☐ Name ☒ Identifier ☐ Reference ☐ NM type  Search

Nanomaterials Advanced search Download

Showing from 1 to 10 in pages of 10 substances Previous Next

JRC

	Substance Name	Substance UUID	Substance Type	Public name	Reference substance UUID	Owner	Info
- 4 -	<a href="#">TiO2 50-150 nm</a>	<a href="#">NNRG-18280a4...</a>	NPO_1486	JRCNM01000a(NM-100)		NANoREG	Material code = JRCNM01000a(NM-100) NANoREG material = Core material Supplier = JRC - IHCP, Fraunhofer
- 5 -	<a href="#">TiO2 6 nm</a>	<a href="#">NNRG-a51b2e5...</a>	NPO_1486	JRCNM01001a (NM-101)		NANoREG	Material code = JRCNM01001a (NM-101) NANoREG material = Core material Supplier = JRC - IHCP, Fraunhofer
- 6 -	<a href="#">TiO2 21-22 nm</a>	<a href="#">NNRG-0bddde2...</a>	NPO_1486	JRCNM01002a(NM-102)		NANoREG	Material code = JRCNM01002a(NM-102) NANoREG material = Core material Supplier = JRC - IHCP
- 7 -	<a href="#">TiO2 24.7 nm</a>	<a href="#">NNRG-818defe7...</a>	NPO_1486	JRCNM01003a(NM-103)		NANoREG	Material code = JRCNM01003a(NM-103) NANoREG material = Core material Supplier = JRC - IHCP
- 8 -	<a href="#">TiO2 23.4 nm</a>	<a href="#">NNRG-91ca30a4...</a>	NPO_1486	JRCNM01005a (NM-105)		NANoREG	Material code = JRCNM01005a (NM-105) NANoREG material = Alternative material Supplier = JRC - IHCP
- 9 -	<a href="#">ZnO 147 nm</a>	<a href="#">NNRG-c6e82a0b...</a>	NPO_1542	JRCNM01100a (NM-110)		NANoREG	Material code = JRCNM01100a (NM-110) NANoREG material = Core material Supplier = JRC - IHCP

Composition name: ZnO Zincite  
Composition UUID: NNRG-c6e82a0b-1eac-da0b-46ad-d2bcb994d1ac  
Purity of IUC Substance:

Type	Name	EC No.	CAS No.	Typical concentration	Concentration ranges	Also contained	Structure
Core				0 % (w/w)	0 % (w/w)	0 % (w/w)	Also contained in...

Chemical structure: O=[Zn]

<http://www.nanoreg.eu>



## Installed:

- A separate instance of the database
- Search application

## Data sources:

- TNO SQL database (converted into eNM SQL)
- Excel files (mapping ongoing)





# NANoREG example: phys-chem

▼ P-Chem

- ☐ 4.27. Nanomaterial aspect ratio/shape (S) [15]
- ☒ 4.28. Nanomaterial specific surface area (S) [111]
- ☐ 4.29. Nanomaterial zeta potential (S) [87]
- ☐ 4.30. Nanomaterial surface chemistry (S) [38]
- ☐ 4.31. Nanomaterial dustiness (S) [3]
- ☐ 4.33. Nanomaterial pour density (S) [4]
- ☐ 4.5. Particle size distribution (Granulometry) (S) [1868]
- ☐ 4.8. Water solubility (S) [155]
- ☐ 4.99. Physico chemical properties (other) (S) [185]

95 materials  
~8475 data points

IUC Substance Composition Tox (81) P-Chem (44)

Filter...

Expand all Collapse all

JRCNM01100a (NM-110)

4.5 Particle size distribution (Granulometry) (25)

4.28 Nanomaterial specific surface area (2)

## Specific Surface Area

Reference	Protocol	Endpoint	Result	Owner	Reliability
JRCNM0110a (2016)	VSSA  T.absorbentgas: Nitrogen T.analysispoints: 83 undefined T.degassingramp: Yes T.endrelativepressure: 0.997 undefined T.instrumentmanufacturer: Micromeritics T.instrumentmodel: TriStar II T.instrumenttype: N2 Sorption T.outgassingtemp: 30 Celsius T.outgassingtime: 6 hour T.sampleweight: 73 mg T.startrelativepressure: 0.002 undefined Dispersion protocol: None Vial: JRCNM01100a	SBET	6.8 m <sup>2</sup> /g	NRCWE	
		TOTAL SURFACE ST	6.8 m <sup>2</sup> /g		
		EXTERNAL SURFACE ST	2.2 m <sup>2</sup> /g		
		MICROPOROSITY SURFACE SU	4.5 m <sup>2</sup> /g		
		Provided	Supplier		

Showing 2 study(s) (1 to 2) Previous Next

4.29 Nanomaterial zeta potential (1)

4.30 Nanomaterial surface chemistry (1)

4.99 Physico chemical properties (other) (15)



# NANoREG example: bioassay

▼ Tox

- ☐ 4.99. Physico chemical properties (other) (S) [1]
- ☐ 7.5.1. Repeated dose toxicity - oral (S) [39]
- ☐ 7.99. Toxicity (other) (S) [169]
- ☐ BAO\_0002167. Genotoxicity Assay (S) [571]
- ☐ BAO\_0002168. Oxidative Stress Assay (S) [10]
- ☒ BAO\_0003009. Cell Viability Assay (S) [846]
- ☐ ENM\_0000044. Barrier integrity (S) [65]
- ☐ NPO\_1339. Immunotoxicity (S) [214]

IUC Substance Composition Tox (81) P-Chem (44)

Filter...

Expand all Collapse all

JRCNM01100a (NM-110)

NPO\_1339 Immunotoxicity (2)

ENM\_0000044 Barrier integrity (4)

BAO\_0002167 Genotoxicity Assay (31)

BAO\_0003009 Cell Viability Assay (44)

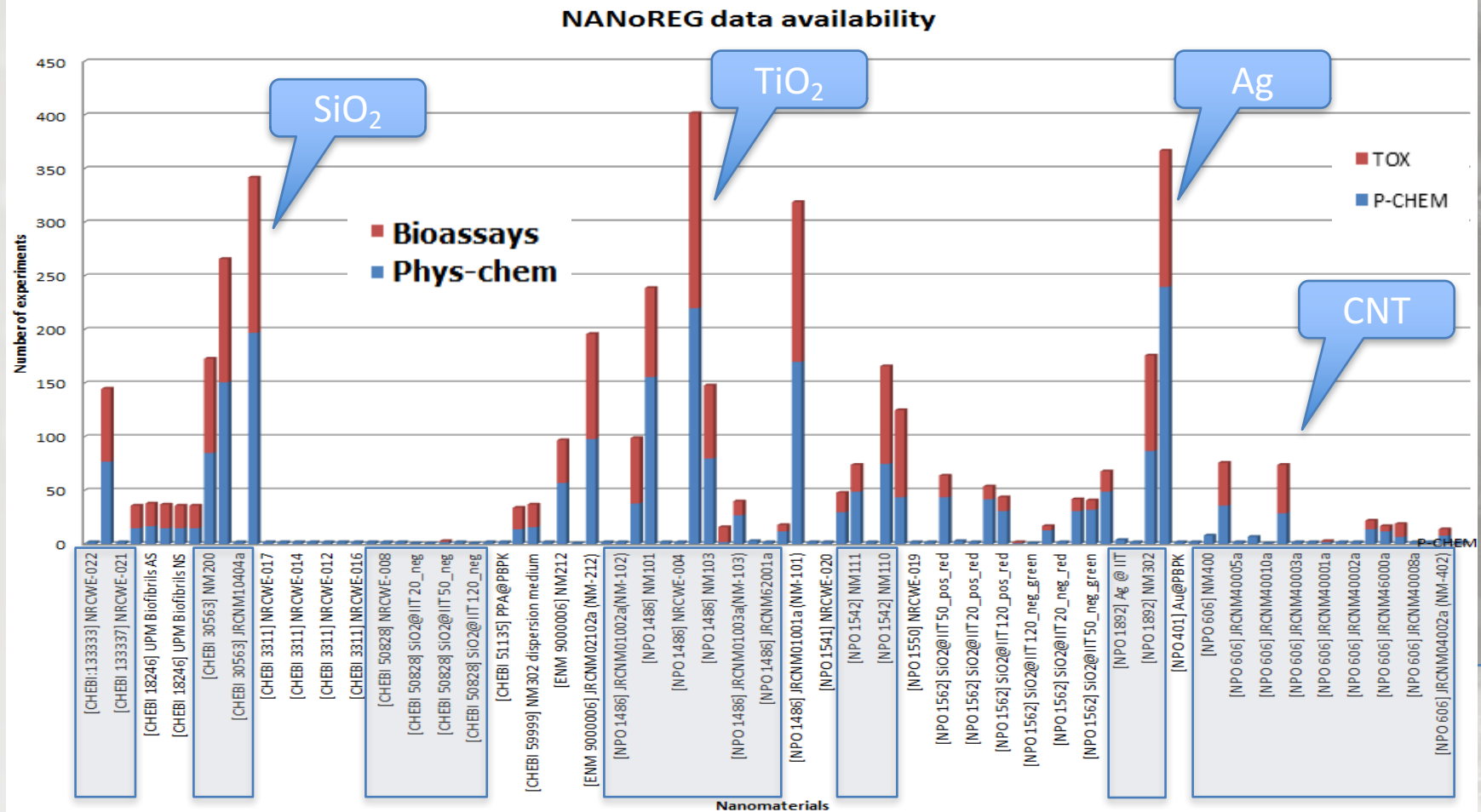
**Cell viability**

Reference	Protocol	Endpoint	Result	concentration	Owner/Reliability
A549 NM110 non-sonicated run 1 48h (2015)	MTS	MAXIMUM TEER VALUES BEFORE STARTING NPS EXPOS	23.255 ohm/cm2	aggregated	FUNDP
	MEDIUM: MEM with glutamax + 10% FBS (A549 cells)	PERCENTAGE VIABILITY COMPARED TO CONTROL	17.203 %	0.000 ug/ml	
	MEDIUM.temperature: 37 Celsius	PERCENTAGE VIABILITY COMPARED TO CONTROL	100 %	1.000 ug/ml	
	MEDIUM.ionic_strength: 0 m	PERCENTAGE VIABILITY COMPARED TO CONTROL	101.934 %	10.000 ug/ml	
	MEDIUM.CO2_concentration: 5 %v/v	PERCENTAGE VIABILITY COMPARED TO CONTROL	101.284 %	25.000 ug/ml	
	MEDIUM.O2_concentration: 0 %v/v	PERCENTAGE VIABILITY COMPARED TO CONTROL	98.326 %	50.000 ug/ml	
	Dispersion protocol: Dispersion by stirring (UNamur)	PERCENTAGE VIABILITY COMPARED TO CONTROL	99.61 %	100.000 ug/ml	
	Vial: JRCNM01100a020064	PERCENTAGE VIABILITY COMPARED TO CONTROL	83.418 %		
	E.cell_type: A549				
	E.days_of_differentiation: 0 d				
	MTS	MAXIMUM TEER VALUES BEFORE STARTING NPS EXPOS	23.255 ohm/cm2	aggregated	
	MEDIUM: MEM with glutamax + 10% FBS (A549 cells)	PERCENTAGE VIABILITY COMPARED TO CONTROL	42.123 %	0.000 ug/ml	
	MEDIUM.temperature: 37 Celsius	PERCENTAGE VIABILITY COMPARED TO CONTROL	100 %		
	MEDIUM.ionic_strength: 0 m				

95 materials  
~16876 data points



# NANoREG data availability (as of Oct 2016)



# NANoREG data availability (as of Oct 2016)

	BaO4S		SiO2	CCaO3.	graphite		Medium	CeO2		TiO2		ZnO		Ag	Au	CNT						
Sum of num	Column																					
Row Labels		CHEBI 133	CHEBI 133	CHEBI 182	CHEBI 305	CHEBI 331	CHEBI 334	CHEBI 508	CHEBI 511	CHEBI 595	CHEBI:133	ENM 900K	ENM 900K	NPO 1486	NPO 1541	NPO 1542	NPO 1550	NPO 1562	NPO 1892	NPO 401	NPO 606	Grand Total
P-CHEM		77	2	77	435	14	2	11	4	30	2	158	2	718	2	200	4	251	333	2	142	2466
ASPECT_RATIO_SHAPE				2					2		2			4				5				15
DUSTINESS														1							2	3
PC_GRANULOMETRY		63	1	77	359	7	1	10	2	27	1	109	1	501	1	156	2	250	204	1	95	1868
DLS		59		71	341					24		94		456		148		229	187		73	1682
SAXS								3										8	1			12
Supplier		3	1	6	7	7	1	7	2		1	6	1	18	1	6	2	13	6	1	21	110
TEM				11					3		9			22		1			9		1	56
WAXD		1												2								3
WAXS														3		1			1			5
PC_UNKNOWN		2			38						21			73		16			18		17	185
PC_WATER_SOL								1						64					90			155
POUR_DENSITY					2									2								4
SPECIFIC_SURFACE_AREA		4	1		14	7	1	1	2		1	6	1	30	1	10	2	1	1	1	27	111
SURFACE_CHEMISTRY		2			7							9		10		9			1			38
ZETA_POTENTIAL		6			13							11		33		9			14		1	87
TOX		68		106	348			1		41		138		576		215		89	216		117	1915
BAO_0002167		24		10	106					15		42		161		58		1	73		81	571
BAO_0002168		2										2		6								10
BAO_0003009		34		35	144					22		44		235		128		85	96		23	846
Alamar Blue		7		10	20					6		10		38		19			14		7	131
Impedance adherent cells		11			47					16		12		105		56		72	35			354
Impedance flow cytometry					11									12				9	7			39
LDH		2		15	5							3		22								47
MTS		12			61							17		52		53		4	40		16	255
Resazurin		2										2		6								10
Trypan Blue				10																		10
ENM_0000044					18							6		24		8		3	3		3	65
NPO_1339		5		31	46			1		4		6		46		21			44		10	214
PC_UNKNOWN														1								1
TO_REPEATED_ORAL					34									5								39
UNKNOWN_TOXICITY		3		30								38		98								169
Grand Total		145	2	183	783	14	2	12	4	71	2	296	2	1294	2	415	4	340	549	2	259	4381





- Free text / faceted search
- Chemistry structure and similarity search
- Data access via API
- Semantic search
- Search integration

# SEARCH TOOLS

# NANoREG DB search application

<https://sandbox.ideaconsult.net/search/nanoreg1>

Search query

carbon nanotube

NANoREG search page (demo)

Home Search Data collections Data upload For developers Help

Search

carbon nanotube

Hits list Selection (3)

Immunotoxicity Particle size distribution (Granulom... carbon nanotube

Clear

< 1 > displaying 1 to 19 of 19

**JRCNM04000a (NM-400) (MWCNT 15.6 nm) carbon nanotube**  
P-CHEM.Particle size distribution (Granulometry) [DLS] [2015]  
[more](#)  
[Material](#) [Composition](#) [Study](#) [External database](#) [Add to Selection](#)

**JRCNM04001a (NM-401) (MWCNT 64.2 nm) carbon nanotube**  
P-CHEM.Particle size distribution (Granulometry) [DLS] [2015]  
[more](#)  
[Material](#) [Composition](#) [Study](#) [External database](#) [Add to Selection](#)

**JRCNM04002a (NM-402) (MWCNT 12.7 nm) carbon nanotube**  
P-CHEM.Particle size distribution (Granulometry) [DLS] [2016]  
[more](#)  
[Material](#) [Composition](#) [Study](#) [External database](#) [Add to Selection](#)

**JRCNM04003a (NM-403) (MWCNT 12 nm) carbon nanotube**  
P-CHEM.Particle size distribution (Granulometry) [DLS] [2015]  
[more](#)  
[Material](#) [Composition](#) [Study](#) [External database](#) [Add to Selection](#)

**JRCNM40001a (MWCNT 15 nm) carbon nanotube**  
P-CHEM.Particle size distribution (Granulometry) [SUPPLIER] [0]  
[more](#)  
[Material](#) [Composition](#) [Study](#) [External database](#) [Add to Selection](#)

**JRCNM40002a (MWCNT 15 nm) carbon nanotube**  
P-CHEM.Particle size distribution (Granulometry) [SUPPLIER] [0]  
[more](#)  
[Material](#) [Composition](#) [Study](#) [External database](#) [Add to Selection](#)

Selected filters

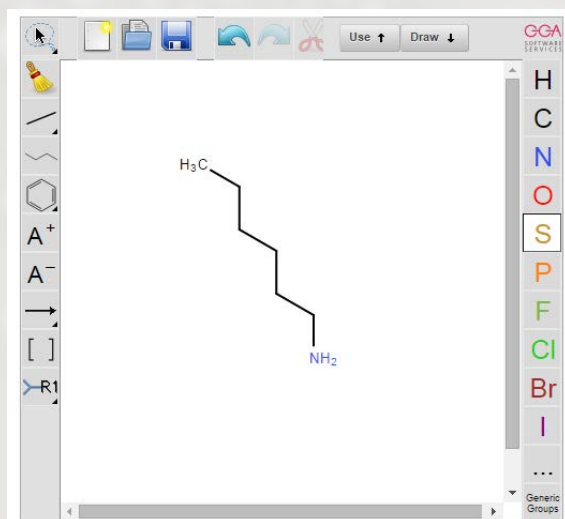
List of data resources with direct links to DB

Facets or filters that permit easy refinement of search

- Data sources (19)
- Nanomaterial type (92)
- P-CHEM (18)
  - Particle size distribution (Granulometry) +
  - Specific surface area +
- TOX (1)
- Medium (1)
- Dispersion protocol (1)
- Results
- References (19)
- Protocols (19)
- Instruments (0)

# Search tools: chemistry

- Chemical structure search: exact, similarity, substructure



Chemical similarity is a pivotal concept in cheminformatics, encompassing a variety of computational methods quantifying the extent to which two chemical structures resemble each other.

Search structures and associated data

Exact structure Similarity Substructure URL

Filter by substance 0.5 C(N)CCCCC

Identifiers Datasets Export

Showing from 1 to 3 in pages of 20 entries Previous Next

Diagram	CasRN	EC number	IUCLID 5 R	Names	Trade Name	IUPAC name	SMILES	Std. InChI key	Std. InChI	REACH registration date	Similarity
- 1 - 			FCSV-97...	Octadecylamine, octadecan-1-amine REYJJPVUYRZGE-UHFFFAOYSA-N InChI=1S/C18H39N/c1-2-3-4-5-6-7-8-9-10-11-12-13-14-15-16-17-18-19/h2-19H2,1H3 stearylamine 1-octadecanamine	-	-	NCCCCC...	REYJJPVUYRZGE-UHFFFAOYSA-N	InChI=1S...	-	0.72
- 2 - 											
- 3 - 											

Showing from 1 to 4 in pages of 20 substances Previous Next

Substance Name	Substance UUID	Substance Type	Public name	Reference substance UUID	Owner	Info	Contained in as
- 1 - 	S40.HDA	FCSV-0e...	NPO_1892	S40.HDA	FCSV-9a...	Protein Corona Fingerprinting Predicts the Cellular Interaction of Gold and Silver Nanoparticles.csv	Classification = Cationic coating 0
- 2 - 	G60.HDA	FCSV-ac...	NPO_401	G60.HDA	FCSV-50...	Protein Corona Fingerprinting Predicts the Cellular Interaction of Gold and Silver Nanoparticles.csv	Classification = Cationic coating 0
- 3 - 	G15.HDA	FCSV-cdf...	NPO_401	G15.HDA	FCSV-50...	Protein Corona Fingerprinting Predicts the Cellular Interaction of Gold and Silver Nanoparticles.csv	Classification = Cationic coating 0
- 4 - 	G30.DDT@HDA	FCSV-cb...	NPO_401	G30.DDT@HDA	FCSV-50...	Protein Corona Fingerprinting Predicts the Cellular Interaction of Gold and Silver Nanoparticles.csv	Classification = Cationic coating 0

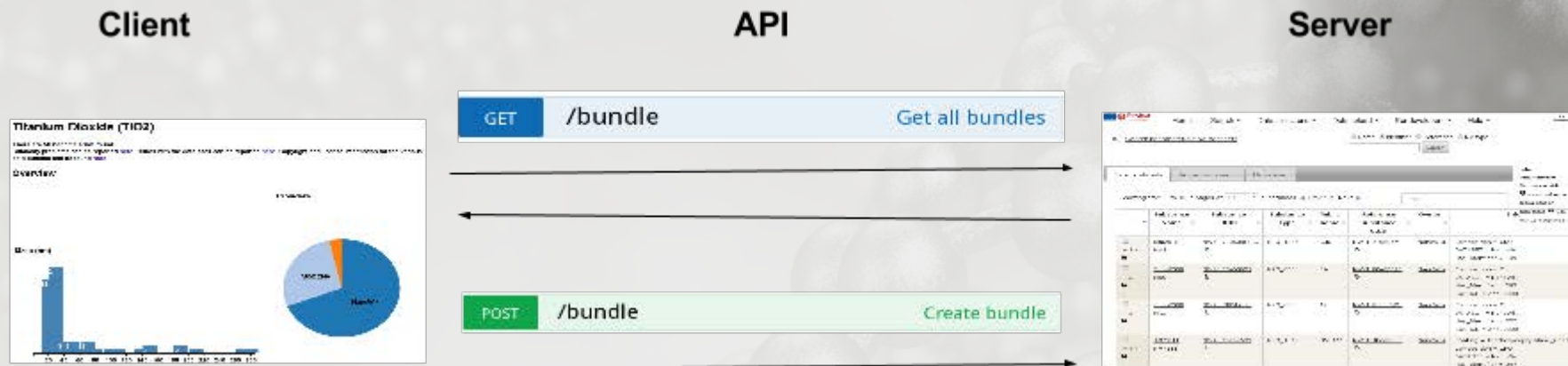
Coating





# Data access: web browser, API

## <http://enanomapper.github.io/API/>



- *REST 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>



# Search data integration:

## <https://search.data.enanomapper.net>

The screenshot displays the eNanoMapper search interface. The main search bar contains the term 'nanotube'. The results are listed under the 'Hits list' tab, showing 1 to 20 of 154 results. The first three results are highlighted with blue arrows pointing to their respective data source details.

**Search Results:**

- Multi-Walled Carbon Nanotubes (MWCNT), synthetic shape mono constituent substance [12.7nm]**  
P-CHEM. Specific surface area [DIN66131] [more](#)  
Material Composition Study [Add to Selection](#)
- STANFORD-ZLiuCR2008-02 carbon nanotube**  
P-CHEM. Particle size distribution (Granulometry) The average length of SWNT-PTX was 106 nm with a standard deviation of 64 nm. [size] [2008] [more](#)  
[caNanoLab](#)
- STANFORD-ZLiuCR2008-01 carbon nanotube**  
P-CHEM. Particle size distribution (Granulometry) Dynamic light scattering hydrodynamic size of SWNTs before and after PTX conjugation of 120. respectively [size] [2008] [more](#)  
[caNanoLab](#)
- SNU-NJiaNL2007-01 multi-walled nanotube**  
P-CHEM. Particle size distribution (Granulometry) [2007] [more](#)  
[caNanoLab](#)

**Data sources (140):**

- NanoWiki, DNT, Nanotechnology ...
- SNU\_CHINA, STANFORD
- STANFORD\_Che..., TAM\_UT
- UM, UNC

**Nanomaterial type (153):**

- P-CHEM (65)
- TOX (32)
- Cell (20)
- Species (1)
- Results (0)
- References (74)
- Protocols (44)
- Instruments (0)

**caNanoLab Details:**

Navigation Tree: HOME, PROTOCOLS, SAMPLES, PUBLICATIONS, HELP, GLOSSARY, LOGIN

Sample: STANFORD-ZLiuCR2008-02

General Info: Sample Name (STANFORD-ZLiuCR2008-02), Created Date (1/4/11), Keywords (EX VIVO IN VIVO THERAPEUTIC EFFICACY TUMOR)

Point of Contact: Primary Contact? (Yes), Contact Person (Hongjie Dai), Organization (STANFORD\_ChemID Department of Chemistry, Stanford University, Stanford CA 94305 USA), Role (Investigator)



- ISA-TAB, ISA-TAB-NANO
- ISA-JSON
- Excel spreadsheets
- Export formats

## I/O SUPPORT: ISA & EXCEL TEMPLATES

# ISA-TAB/ISA-JSON



## Version 1 – ISA-TAB (Nov 2008)

Data is described in 3 layers

Tab delimited format (\*.txt)

Only meta data is stored

Pointers to the real data

Ontology references

Additional configurations

## ISA-JSON version 1

(officially released 2016)

## ISA-JSON Version 2

(under development)

30 November 2016



### investigation

high level concept to link related studies

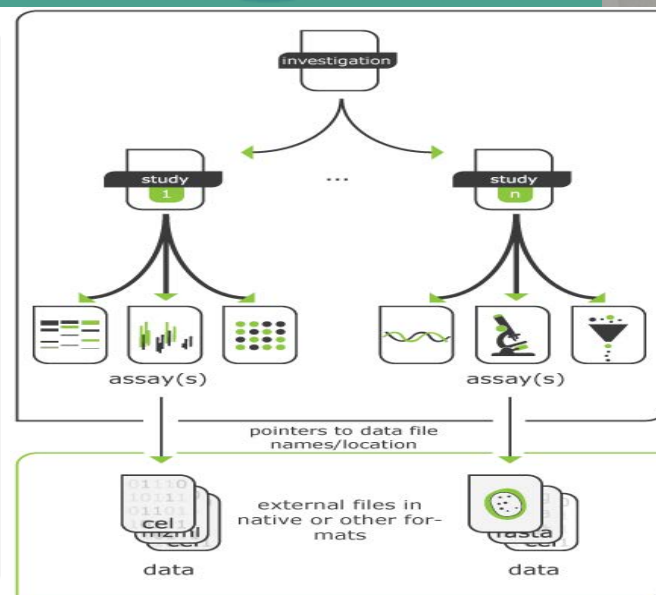
### study

the central unit, containing information on the subject under study, its characteristics and any treatments applied.

*a study has associated **assays***

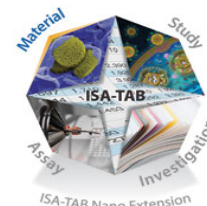
### assay

test performed either on material taken from the subject or on the whole initial subject, which produce qualitative or quantitative measurements (data)

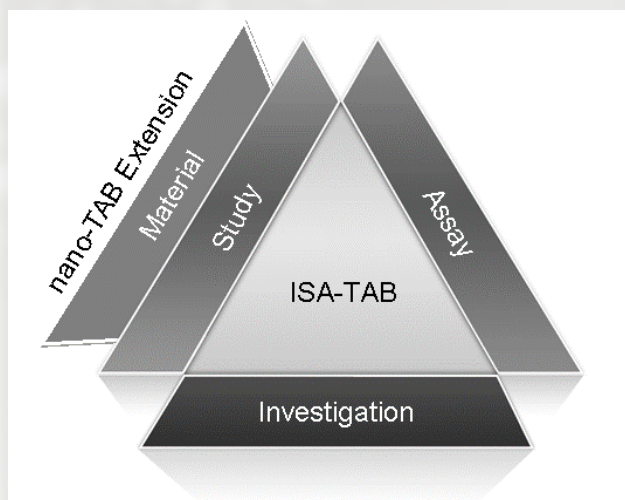




# ISA-TAB-Nano



ISA-TAB-Nano uses the three primary files of ISA-TAB investigation file, study file, and assay file; and introduces a fourth file called the **material file**.

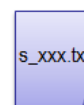


1. Describe the Investigation and Studies



Investigation File

2. Identify Study Samples

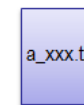


Study File(s)



Material File(s)

3. Record Assay Conditions and Measurements



Assay File(s)

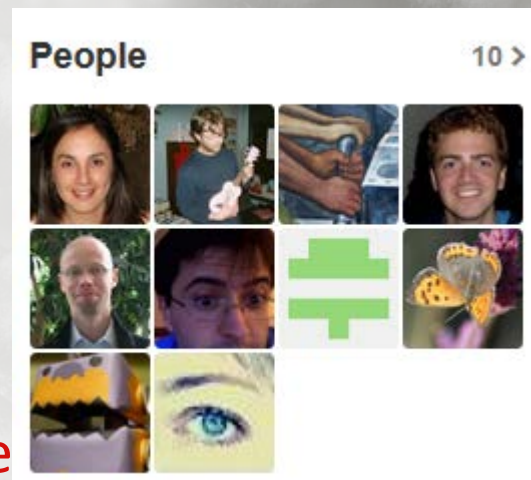


# ISA-JSON project



<https://github.com/ISA-tools/isa-api>

- Developed by S.Sansone group (University of Oxford) and collaborators
- Python based **ISA** API library
- New data format based on JSON describes the ISA experimental graph
- Full support of ISA-TAB (released ISA-JSON v.1)
- More efficient data storage than the TAB delimited
- New extended ISA v.2 (under development)



# ISA-JSON schemas



[https://github.com/ISA-tools/isa-api/tree/master/isatools/schemas/isa\\_model\\_version\\_1\\_0\\_schemas/core](https://github.com/ISA-tools/isa-api/tree/master/isatools/schemas/isa_model_version_1_0_schemas/core)

assay\_schema.json

comment\_schema.json

data\_schema.json

factor\_schema.json

factor\_value\_schema.json

investigation\_schema.json

material\_attribute\_schema.json

material\_schema.json

ontology\_annotation\_schema.json

ontology\_source\_reference\_schema.json

organization\_schema.json

parameter\_schema.json

person\_schema.json

process\_schema.json

protocol\_schema.json

```
{
  "$schema": "http://json-schema.org/draft-04/schema",
  "title": "ISA investigation schema",
  "description": "JSON-schema representing an investigation in the ISA model",
  "type": "object",
  "properties": {
    "identifier": { "type": "string" },
    "title": { "type": "string" },
    "description": { "type": "string" },
    "submissionDate": { "type": "string", "format": "date-time" },
    "publicReleaseDate": { "type": "string", "format": "date-time" },
    "commentCreatedWithConfiguration": {
      "$ref": "comment_schema.json#",
      "name": "Created With Configuration"
    },
    "commentLastOpenedWithConfiguration": {
      "$ref": "comment_schema.json#",
      "name": "Last Opened With Configuration"
    },
    "ontologySourceReferences": {
      "type": "array",
      "items": {
        "$ref": "ontology_source_reference_schema.json#"
      }
    }
  }
}
```

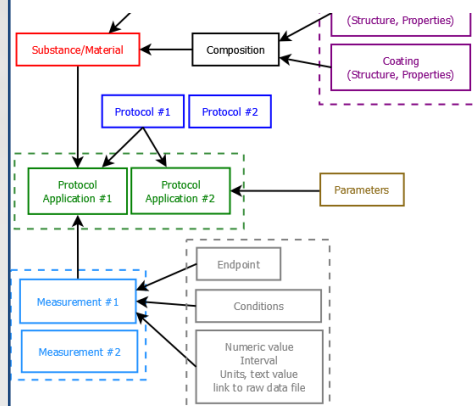
# eNanoMapper ISA-JSON export

## ISA JSON schemas

- assay\_schema.json
- comment\_schema.json
- data\_schema.json
- factor\_schema.json
- factor\_value\_schema.json
- investigation\_schema.json
- material\_attribute\_schema.json
- material\_schema.json
- ontology\_annotation\_schema.json
- ontology\_source\_reference\_schema.json
- organization\_schema.json
- parameter\_schema.json
- person\_schema.json
- process\_schema.json
- protocol\_schema.json
- publication\_schema.json
- sample\_schema.json
- source\_schema.json
- study\_schema.json

**Material JSON  
schema**

## eNM data model



**Conversion  
to pojo**

**Store eNM substances as  
objects corresponding to the  
ISA schemas**

- Assay.java
- Comment.java
- Component.java
- Data.java
- Factor.java
- FactorValue.java
- Investigation.java
- Material.java
- MaterialAttribute.java
- MeasurementType.java
- OntologyAnnotation.java
- OntologySourceReference.java

**Export  
configuration**

*additional data files*

**Single  
ISA-JSON file**

**export**

**isa-api**

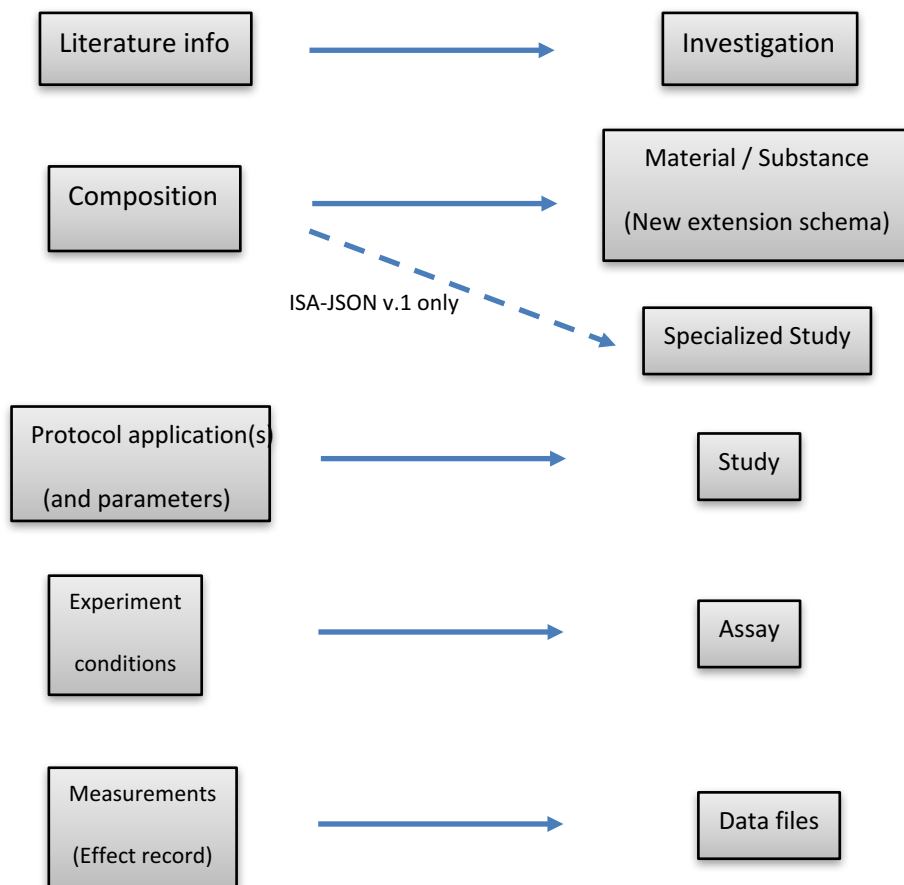
**ISA-TAB files**



# eNanoMapper – ISA mapping

## ENM data model

## ISA platform



## Workflow

- The ISA JSON schema is used to generate Java classes (*next slide*)
- The ISA Java classes correspondence to the eNanoMapper data model (*this slide*)
- The data is loaded into eNanoMapper data model
- Converted into ISA model
- Exported into ISA-JSON
- ISA-JSON can be converted to ISA-TAB



# ISA (v.1) Java classes

- Assay.java
- Comment.java
- Component.java
- Data.java
- Factor.java
- FactorValue.java
- Investigation.java
- Material.java
- MaterialAttribute.java
- MeasurementType.java
- OntologyAnnotation.java
- OntologySourceReference.java
- Organization.java
- Parameter.java
- Person.java
- Process.java
- Protocol.java
- Publication.java
- Role.java
- Sample.java
- Source.java
- Study.java
- TechnologyType.java
- Value.java

Investigation.java

- Investigation
  - commentCreatedWithConfiguration
  - commentLastOpenedWithConfiguration
  - description
  - identifier
  - ontologySourceReferences
  - people
  - publications
  - publicReleaseDate
  - studies
  - submissionDate
  - title

```
@JsonProperty("identifier")
public String identifier;
@JsonProperty("title")
public String title;
@JsonProperty("description")
public String description;
@JsonProperty("submissionDate")
public Date submissionDate;
@JsonProperty("publicReleaseDate")
public Date publicReleaseDate;

@JsonProperty("commentCreatedWithConfiguration")
public Comment commentCreatedWithConfiguration;
@JsonProperty("commentLastOpenedWithConfiguration")
public Comment commentLastOpenedWithConfiguration;
@JsonProperty("ontologySourceReferences")
public List<OntologySourceReference> ontologySourceReferences = new ArrayList<OntologySourceReference>();
@JsonProperty("publications")
public List<Publication> publications = new ArrayList<Publication>();
@JsonProperty("people")
public List<Person> people = new ArrayList<Person>();
@JsonProperty("studies")
public List<Study> studies = new ArrayList<Study>();
}
```

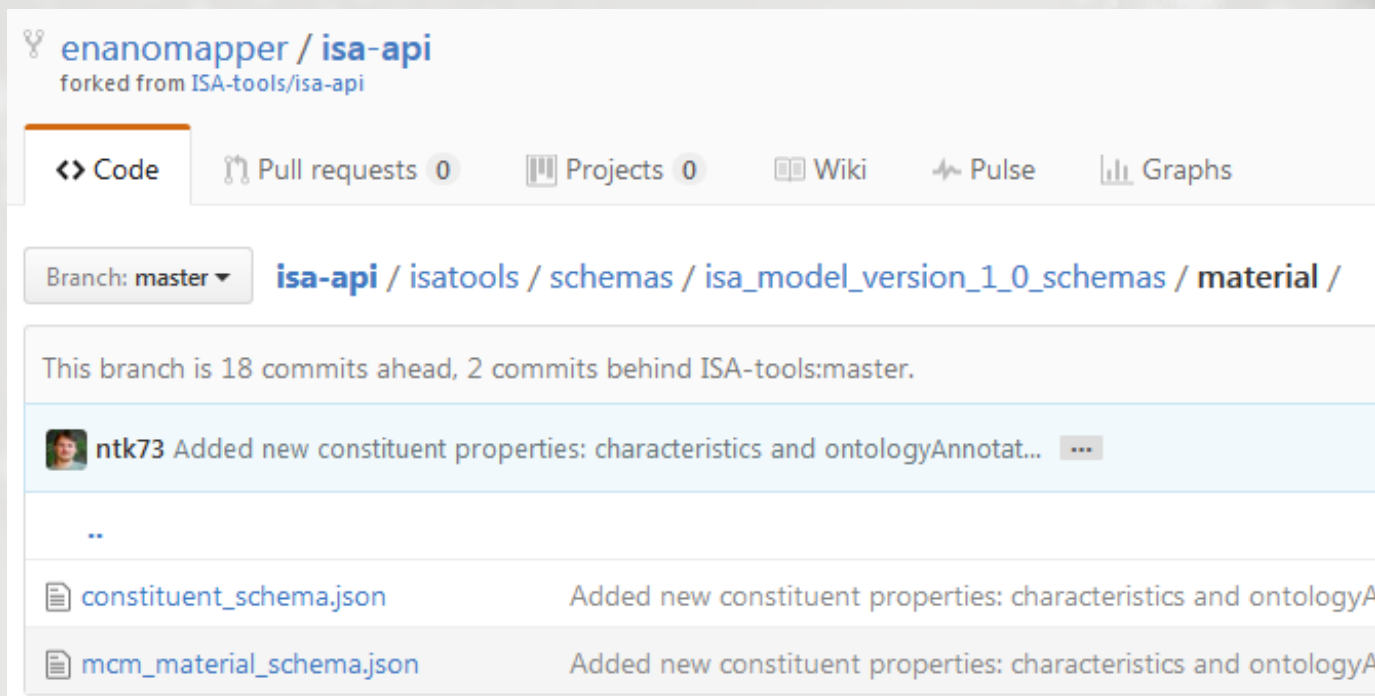
Study.java

- Study
  - assays
  - description
  - identifier
  - people
  - processSequence
  - protocols
  - publications
  - publicReleaseDate
  - samples
  - sources
  - studyDesignDescriptors
  - submissionDate
  - title

```
public class Study {

    @JsonProperty("identifier")
    public String identifier;
    @JsonProperty("title")
    public String title;
    @JsonProperty("description")
    public String description;
    @JsonProperty("submissionDate")
    public Date submissionDate;
    @JsonProperty("publicReleaseDate")
    public Date publicReleaseDate;
    @JsonProperty("publications")
    public List<Publication> publications = new ArrayList<Publication>();
    @JsonProperty("people")
    public List<Person> people = new ArrayList<Person>();
    @JsonProperty("studyDesignDescriptors")
    public List<OntologyAnnotation> studyDesignDescriptors = new ArrayList<OntologyAnnotation>();
    @JsonProperty("protocols")
    public List<Protocol> protocols = new ArrayList<Protocol>();
}
```

# ISA-JSON material extension



enanomapper / isa-api  
forked from ISA-tools/isa-api

Code Pull requests 0 Projects 0 Wiki Pulse Graphs

Branch: master isa-api / isatools / schemas / isa\_model\_version\_1\_0\_schemas / material /

This branch is 18 commits ahead, 2 commits behind ISA-tools:master.

ntk73 Added new constituent properties: characteristics and ontologyAnnotat... ..

..

constituent\_schema.json Added new constituent properties: characteristics and ontologyA

mcm\_material\_schema.json Added new constituent properties: characteristics and ontologyA

**Contributing new  
extension to isa-api**

**JSON schema corresponding  
to ISA-TAB-Nano  
material file**

# ISA-JSON material schema

```
$schema : http://json-schema.org/draft-04/schema#  
title : Material  
description : Definition of Material (or Substance)  
type : object
```

## ▼ properties {14}

- ▶ @id {2}
- ▶ lotIdentifier {1}
- ▶ name {2}
- ▶ sourceName {1}
- ▶ root {2}
- ▶ **constituents {2}**
- ▶ description {1}
- ▶ synthesis {1}
- ▶ designRationale {1}
- ▶ intendedApplication {2}
- ▶ characteristics {2}
- ▶ mcmType {2}
- ▶ chemicalName {2}
- ▶ dataFiles {2}

```
{  
  "$schema": "http://json-schema.org/draft-04/schema#",  
  "title": "Constituent",  
  "description": "Definition of a constituent of a material or another  
    constituent",  
  "type": "object",  
  "properties": {  
    "@id": { "type": "string", "format": "uri" },  
    "name": {  
      "type": "string",  
      "description": "Constituent name"  
    },  
    "role": { "type": "string" },  
    "description": { "type": "string" },  
    "synthesis": { "type": "string" },  
    "linkages": {  
      "type": "array",  
      "items": {  
        "type": "object",  
        "properties": {  
          "constituent": { "type": "string", "format": "uri" },  
          "linkageType": { "type": "string" }  
        }  
      }  
    },  
    "characteristics" : {  
      "type" : "array",  
      "items" : {  
        "$ref": "material_attribute_value_schema.json#"  
      }  
    },  
    "ontologyAnnotation" : {  
      "$ref": "ontology_annotation_schema.json#"  
    }  
  }  
}
```





# Data export: ISA-JSON, RDF, etc.

eNanoMapper

https://data.enanomapper.net/bundle

Home Search Data collections Data upload For developers Help Log in

Bundles

Name Search

Nanomaterial bundles Download Help

Showing 9 bundles (1 to 9)

Bundle ID	Name	Ver	Description	Source	License & Maintainer
00000000-0000-0000-0000-000000000001	<b>NanoWiki</b>	v2	Nanomaterials, physicochemical characterisations and toxicity data, imported via NanoWiki RDF dump	NanoWiki <a href="http://dx.doi.org/10.6084/m9.figshare.1330208">http://dx.doi.org/10.6084/m9.figshare.1330208</a>	<a href="http://orcid.org/0000-0001-7542-0286">http://orcid.org/0000-0001-7542-0286</a>
<b>Details</b>					
Substances (389)		<a href="#">Browse substances</a>		<b>Download as</b>	
Chemical structures (74)		<a href="#">Browse chemical structures</a>			
Dataset		<a href="#">Browse dataset</a>			
Properties (7)		<a href="#">Browse properties</a>			
00000000-0000-0000-0000-000000000002	<b>OECD Harmonized Templates import test</b>	v1	Demonstrate import of a nanomaterial and endpoints data form IUCLID5 *.i5z file (OECD HT)	Multi-Walled Carbon Nanotubes (MWCNT), synthetic graphite in tubular shape <a href="http://apps.echa.europa.eu/registered/data/dossiers/DISS-b281d1a0-c6d8-5dcf-e044-00144f67d031/AGGR-cd35254a-7b90-4a1f-842d-7700c6a210e9_DISS-b281d1a0-c6d8-5dcf-e044-00144f67d031.html">http://apps.echa.europa.eu/registered/data/dossiers/DISS-b281d1a0-c6d8-5dcf-e044-00144f67d031/AGGR-cd35254a-7b90-4a1f-842d-7700c6a210e9_DISS-b281d1a0-c6d8-5dcf-e044-00144f67d031.html</a>	<a href="http://www.ideaconsult.net">www.ideaconsult.net</a>
00000000-0000-0000-0000-000000000003	<b>Protein Corona Fingerprinting Predicts the Cellular Interaction of Gold and Silver Nanoparticles</b>	v1	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 <a href="http://pubs.acs.org/doi/abs/10.1021/nn406018q">http://pubs.acs.org/doi/abs/10.1021/nn406018q</a>	NTUA
91cad054-4de8-4dc9-a8e4-20564e7eada7	<b>Comprehensive In Vitro Toxicity Testing of a Panel of Representative Oxide Nanomaterials</b>	v1	10.1371/journal.pone.0127174	<a href="http://pubs.acs.org/doi/abs/10.1021/nn406018q">http://pubs.acs.org/doi/abs/10.1021/nn406018q</a>	

Also a command line application:  
<http://ambit.sourceforge.net/enanomapper/templates/convertor.html>





# Data Import: EU NanoSafety Cluster Excel templates

## Two types of Excel templates:

## 1) ISA-TAB Logic templates (NANoREG)

Not strictly following the ISA-TAB and ISA-TAB-Nano formats, designed around ISA-Tab-logic, i.e. structuring the data in investigation-study-assay related groups.

*One sheet: many materials, one assay, both metadata and data; CC BY-SA 4.0 license*

<http://www.nanoreg.eu/media-and-downloads/templates/269-templates-for-experimental-data-logging>

2) One material, one assay;

*first sheet: metadata; next sheets: raw and processed data (used by several EU projects; many variations, not publicly available)*

## Solution: A configurable Excel Parser for custom spreadsheets

## JSON configuration mapping the Excel layout into the eNanoMapper data model

*(next slide)*

<https://github.com/enanomapper/nmdataparser>

TEST CONDITIONS		Please complete the details below as far as possible for each set of assay results	
<p>Assay Name: <input type="text"/></p> <p>Test Facility: <input type="text"/></p> <p>Test Date: <input type="text"/></p> <p>Test Time: <input type="text"/></p> <p>Test Location: <input type="text"/></p> <p>Test Operator: <input type="text"/></p> <p>Test Assistant: <input type="text"/></p> <p>Test Sample: <input type="text"/></p> <p>Test Sample ID: <input type="text"/></p> <p>Test Sample Description: <input type="text"/></p> <p>Test Sample Source: <input type="text"/></p> <p>Test Sample Storage: <input type="text"/></p> <p>Test Sample Handling: <input type="text"/></p> <p>Test Sample Analysis: <input type="text"/></p> <p>Test Sample Results: <input type="text"/></p> <p>Test Sample Comments: <input type="text"/></p>		<p>Assay Name: <input type="text"/></p> <p>Test Facility: <input type="text"/></p> <p>Test Date: <input type="text"/></p> <p>Test Time: <input type="text"/></p> <p>Test Location: <input type="text"/></p> <p>Test Operator: <input type="text"/></p> <p>Test Assistant: <input type="text"/></p> <p>Test Sample: <input type="text"/></p> <p>Test Sample ID: <input type="text"/></p> <p>Test Sample Description: <input type="text"/></p> <p>Test Sample Source: <input type="text"/></p> <p>Test Sample Storage: <input type="text"/></p> <p>Test Sample Handling: <input type="text"/></p> <p>Test Sample Analysis: <input type="text"/></p> <p>Test Sample Results: <input type="text"/></p> <p>Test Sample Comments: <input type="text"/></p>	



# Mapping the spreadsheet content into the data model

	A	B	C	D	E	F	G
1	<b>TEST RESULTS</b>						
2							
3							
4		Replicate 1			Replicate 2		
5		T1	Titanium Dioxide	Average (ng/ml)		T1	Titanium Dioxide
6			0	1.8925			
7			1	19.6985			
8			5	18.5207			
9			10	18.0280			
10			25	18.4153			
11			50	19.2965			
12			75	20.8867			
13			100	22.6964			
14							
15							
16							



through JSON configuration

```
{
  "TEMPL": {
    "DATA_": {
      "PARAL": {
        "SUBSTANCE_RECORD": { ... }, // 7 items
        "PROTOCOL_APPLICATIONS": [
          {
            "PROTOCOL_TOP_CATEGORY": { ... }, // 2 items
            "PROTOCOL_CATEGORY_CODE": { ... }, // 2 items
            "PROTOCOL_GUIDELINE": { ... }, // 1 item
            "PROTOCOL_ENDPOINT": { ... }, // 4 items
            "RELIABILITY_STUDY_RESULT_TYPE": { ... }, // 2 items
            "CITATION_TITLE": { ... }, // 2 items
            "CITATION_YEAR": { ... }, // 2 items
            "CITATION_OWNER": { ... }, // 4 items
            "PARAMETERS": { ... }, // 12 items
            "EFFECTS_BLOCK": {
              "LOCATION": {
                "ITERATION": "ABSOLUTE_LOCATION",
                "IS_ARRAY": true,
                "TRIM_ARRAY": true,
                "SHEET_INDEX": 3,
                "COLUMN_INDEX": "B",
                "ROW_INDEX": 4
              },
              "ROW_SUBBLOCKS": "= TimePoints.size()",
              "COLUMN_SUBBLOCKS": "= Replicates",
              "SUBBLOCK_SIZE_ROWS": "= C.size() + 3",
              "SUBBLOCK_SIZE_COLUMNS": 4,
              "VALUE_GROUPS": [
                {
                  "NAME": "=Assay_endpoint",
                  "UNIT": "=Outcome_metric",
                  "START_COLUMN": 3,
                  "END_COLUMN": 3,
                  "START_ROW": 3,
                  "END_ROW": "=3 - 1 + C.size()",
                  "PARAMETERS": [
                    {
                      "NAME": "Time point",
                      "ASSIGN": "ASSIGN_TO_SUBBLOCK",
                      "COLUMN_POS": 1,
                      "ROW_POS": 2,
                      "MAPPING": "Time",
                      "UNIT": "h"
                    },
                    {
                      "NAME": "Replicate",
                      "ASSIGN": "ASSIGN_TO_SU",
                      "COLUMN_POS": 1,
                      "ROW_POS": 1
                    },
                    {
                      "NAME": "Concentration",
                      "ASSIGN": "ASSIGN_TO_VALUE",
                      "COLUMN_POS": -1,
                      "ROW_POS": 0,
                      "UNIT": "µg/ml"
                    }
                  ]
                }
              ]
            }
          ]
        ]
      }
    }
  }
}
```



← → C <https://apps.ideaconsult.net/enanomapper/ui/uploadsubstance1>

Search ▾ Nanomaterials ▾ OpenTox ▾ Help ▾

⌕ Substances ▸ Import ▸ Single file upload

Import new substance(s)

File (.isz or .isd or .xls or .xlsx )\*  No file chosen

JSON map for XLS/XLSX file  No file chosen




# Automating the JSON configuration (under development)

ambit.sourceforge.net/enanomapper/templates/index.html


<http://ambit.sourceforge.net/enanomapper/templates/>

## eNanoMapper data templates release


### Physicochemical characterisation



### In-vitro assays



### In-vivo assays



### NANoREG templates structure

The NANoREG templates are developed by JRC with collaboration with FP7 NANoREG project partners and released under open license at <http://www.nanoreg.eu/media-and-downloads/templates>.

Within eNanoMapper project (WP3 Database) the templates are analysed, fields cleaned and JSON configuration files created. The number of **unique terms** in the templates is over 800.

The NANoREG templates are organized as one spreadsheet per assay, multiple NM can be entered as rows. One Excel file may contain more than one assay, measuring the same endpoint. The metadata is organised in several groups:

#### About

The Excel templates are automatically regenerated based on fields defined in the **NANoREG templates**. The generation process includes cleaning of the JRC fields and facilitates the automatic generation of **JSON configuration files**, necessary to enable import into an **eNanoMapper database instance**.

While not strictly following the **ISA-TAB** and **ISA-TAB-Nano** formats, the NANoREG templates have been designed around ISA-Tab-logic, i.e. structuring the data in investigation-study-assay related groups.

See more about

- [ISA-TAB and the new ISA-JSON here](#)
- [New \(Nano\)Material schema for ISA-JSON](#)
- [How to convert](#) Excel data files into ISA-JSON or RDF.
- [How to import into eNanoMapper database instance.](#)
- [How the conversion works](#)

UNDER DEVELOPMENT <

- Extract all fields from NANoREG templates;
- Cleanup (typos, units), sync between different templates;
- Annotation;
- Generate the templates based on cleaned fields and JSON configurations;
- **One-assay Excel template + JSON, ready for upload;**
- Next step – dynamic generation



# Finally – a bonus: command line XLSX- ISA-JSON/RDF convertor

[Home](#)[P-CHEM](#)[In-vitro](#)[In-vivo](#)[Usage](#)[Database search](#)

<http://ambit.sourceforge.net/enanomapper/templates/convertor.html>

A command line application for converting between supported data formats with nanomaterial safety data.

- Download <https://github.com/enanomapper/nmdataparser>

Excel spreadsheets import requires a separate JSON configuration file. More details about the **Excel parser**. Predefined JSON configuration files for the **NANoREG templates** are provided next to each Excel file, e.g. **INVITRO/GENOTOXICITY/COMET**.

- Supported formats

```
java -jar enmconvertor.jar -L
```

(RW)	xls	Excel (.xls) spreadsheet, requires JSON configuration file (option -x)
(RW)	xlsx	Excel (.xlsx) spreadsheet, requires JSON configuration file (option -x)
(RW)	json	AMBIT JSON
(RW)	rdf	eNanoMapper RDF (based on BioAssayOntology RDF)
(W)	isa	ISA-JSON v1 (see <a href="https://github.com/ISA-tools/isa-api">https://github.com/ISA-tools/isa-api</a> )
(R)	Nwrd	Nanowiki RDF (Semantic Media Wiki RDF export)

## Examples

### Converting NanoSafety Cluster Excel spreadsheets

- Example `xlsx` file `INVITRO_VIABILITY_Trypanblue_TEST.xlsx`
- Example `json` configuration file `INVITRO_VIABILITY_Trypanblue.json`

to ISA-JSONv1

```
java -jar enmconvertor.jar -i "INVITRO_VIABILITY_Trypanblue_TEST.xlsx" -j "INVITRO_VIABILITY_Trypanblue.json" -I xlsx -O isa -o "INVITRO_VIABILITY_Trypan blue_TEST.isa.json"
```

Result file (zipped) `INVITRO_VIABILITY_Trypanblue_TEST.isa.json.zip`

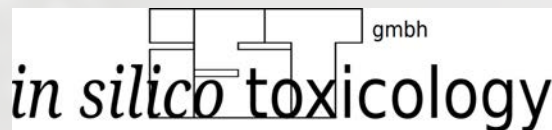




# Summary

- Open source database and web application
- Demo at <https://data.enanomapper.net>
- Import: Excel templates, RDF, OECD OHT, SQL
- Export: ISA-JSON, RDF, XLSX
- Enables distributed setup: many databases; search integration <https://search.data.enanomapper.net>
- Integration with data analysis tools
- Search tools: free text, chemistry, semantic
- More on ontology: NanoWG, Dec 8, by Maastricht U.





Questions?

**THANK YOU!**

