

eNanoMapper database, search tools and templates

Nina Jeliaskova, Nikolay Kochev



IdeaConsult Ltd.

Sofia, Bulgaria

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

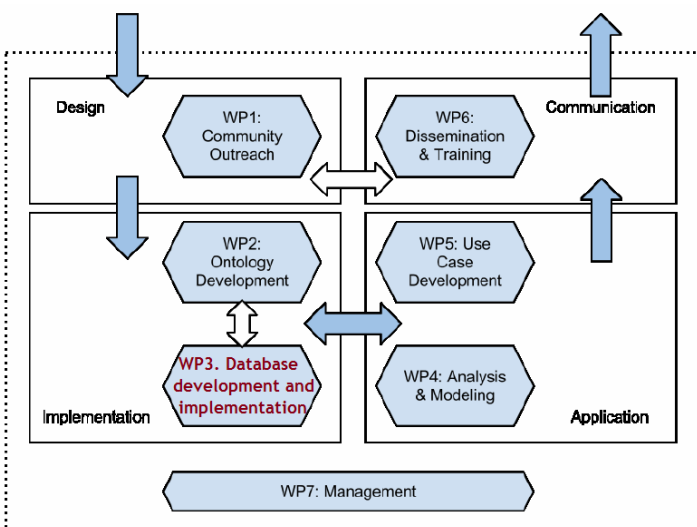
CONTENT



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**

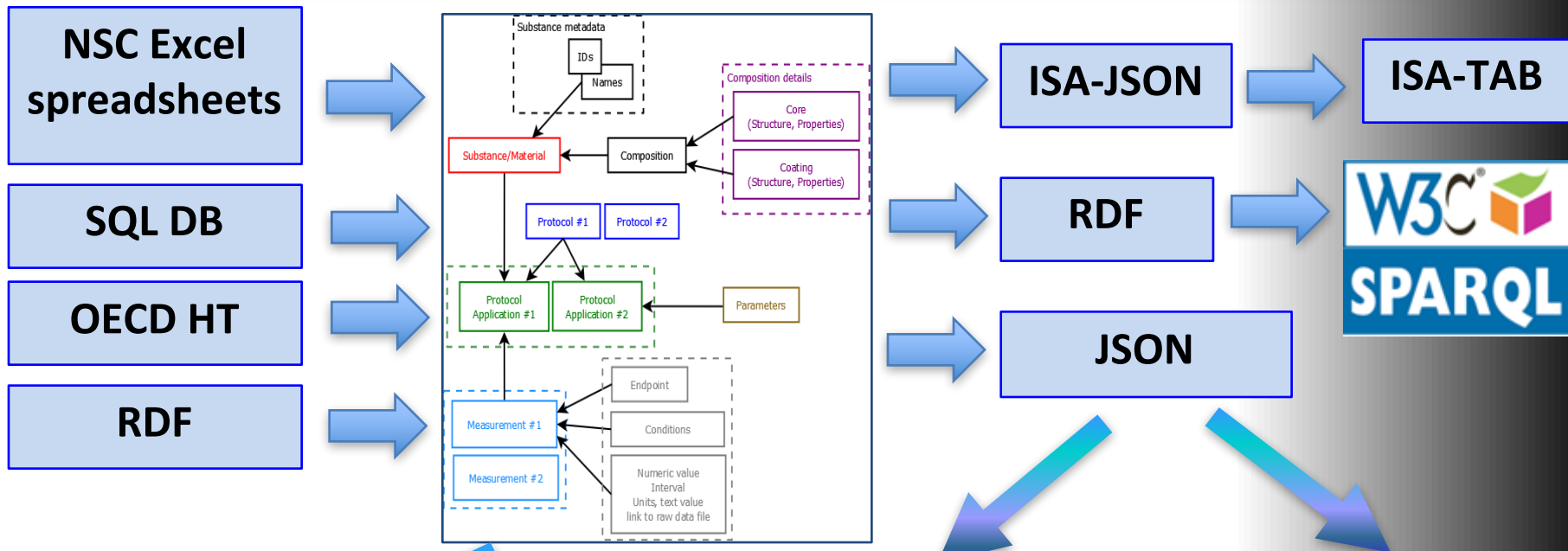
From raw data (science) to study summaries for regulatory purposes; linking with experimental protocols; risk assessment; grouping, safety-by-design

- **Support for data analysis**

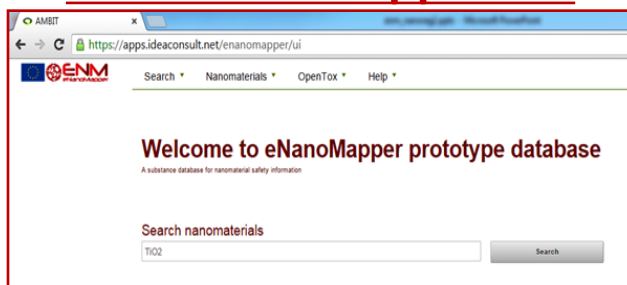
Requires “spreadsheet” or matrix view of data. The experimental data in the public datasets is usually not in a form appropriate for modelling (merging multiple values, conditions, similar experiments into matrix form is a challenge).



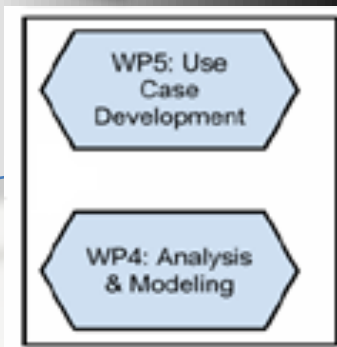
The common eNanoMapper data model : enables conversions



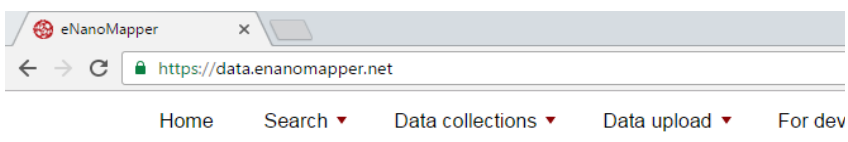
data.enanomapper.net



Yet another search service



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](#) | [by 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](https://doi.org/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)



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 -	TiO2 50-150 nm	NNRG-18280a4...	NPO_1486	JRCNM01000a(NM-100)		NANoREG	Material code = JRCNM01000a(NM-100) NANoREG material = Core material Supplier = JRC - IHCP,Fraunhofer
- 5 -	TiO2 6 nm	NNRG-a51b2e5...	NPO_1486	JRCNM01001a (NM-101)		NANoREG	Material code = JRCNM01001a (NM-101) NANoREG material = Core material Supplier = JRC - IHCP,Fraunhofer
- 6 -	TiO2 21-22 nm	NNRG-0bddde2...	NPO_1486	JRCNM01002a(NM-102)		NANoREG	Material code = JRCNM01002a(NM-102) NANoREG material = Core material Supplier = JRC - IHCP
- 7 -	TiO2 24.7 nm	NNRG-818defe7...	NPO_1486	JRCNM01003a(NM-103)		NANoREG	Material code = JRCNM01003a(NM-103) NANoREG material = Core material Supplier = JRC - IHCP
- 8 -	TiO2 23.4 nm	NNRG-91ca30a4...	NPO_1486	JRCNM01005a (NM-105)		NANoREG	Material code = JRCNM01005a (NM-105) NANoREG material = Alternative material Supplier = JRC - IHCP
- 9 -	ZnO 147 nm	NNRG-c6e82a0b...	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-d2bc994d1ac

Purity of IUC Substance:

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

<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]

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

Filter...

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	SBET	6.8 m ² /g	NRCWE	
	T.absorbentgas: Nitrogen	TOTAL SURFACE ST	6.8 m ² /g		
	T.analysispoints: 83 undefined	EXTERNAL SURFACE ST	2.2 m ² /g		
	T.degassingramp: Yes	MICROPOROSITY SURFACE SU	4.5 m ² /g		
Provided	Supplier	SPECIFIC SURFACE AREA	14 m ² /g	JRC - IHCP	

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)

95 materials
~8475 data points



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...

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

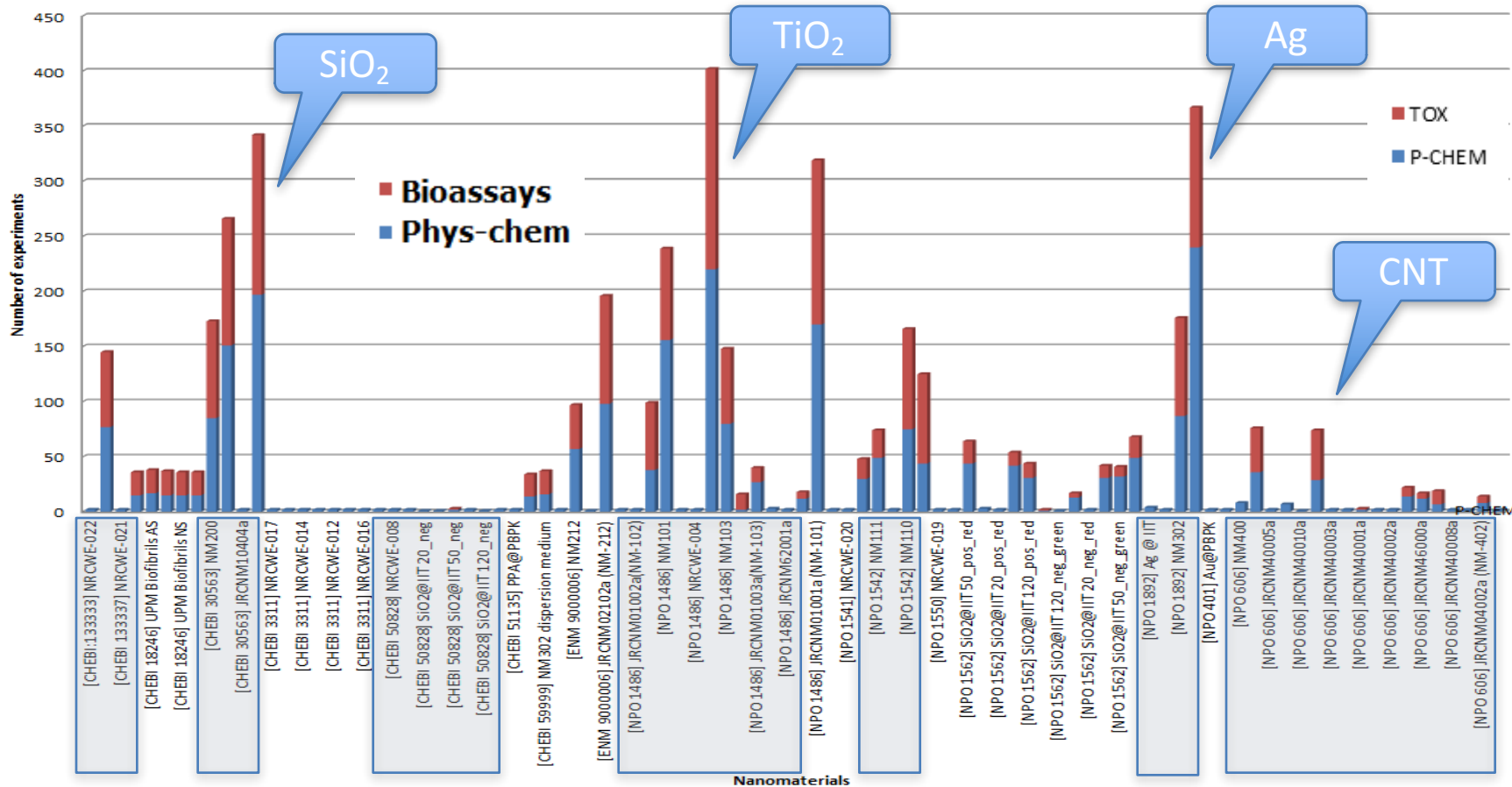
Reference	Protocol	Endpoint	Result	concentration	Owner/Reliability
A549 NM110 non-sonicated run 1 48h (2015)	MTS MEDIUM: MEM with glutamax + 10% FBS (A549 cells) MEDIUM.temperature: 37 Celsius MEDIUM.ionic_strength: 0 m MEDIUM.CO2_concentration: 5 %v/v MEDIUM.O2_concentration: 0 %v/v Dispersion protocol: Dispersion by stirring (UNamur) Vial: JRCNM01100a020064 E.cell_type: A549 E.days_of_differentiation: 0 d E.exposure_time: 48 hour E.method: MTS E.method_supplier: NANOVALID E.positive_controlid: STAUROSPORINE E.sop_reference: NANOVALID SOP FOR MTS ASSAY E.well_plate: 96	MAXIMUM TEER VALUES BEFORE STARTING NPS EXPOS	23.255 ohms/cm2	aggregated	FUNDP
		PERCENTAGE VIABILITY COMPARED TO CONTROL	17.203 %		
		PERCENTAGE VIABILITY COMPARED TO CONTROL	100 %	0.000 ug/ml	
		PERCENTAGE VIABILITY COMPARED TO CONTROL	101.934 %	1.000 ug/ml	
		PERCENTAGE VIABILITY COMPARED TO CONTROL	101.284 %	10.000 ug/ml	
		PERCENTAGE VIABILITY COMPARED TO CONTROL	98.326 %	25.000 ug/ml	
		PERCENTAGE VIABILITY COMPARED TO CONTROL	99.61 %	50.000 ug/ml	
		PERCENTAGE VIABILITY COMPARED TO CONTROL	83.418 %	100.000 ug/ml	
		MAXIMUM TEER VALUES BEFORE STARTING NPS EXPOS	23.255 ohms/cm2	aggregated	
		PERCENTAGE VIABILITY COMPARED TO CONTROL	42.123 %		
		PERCENTAGE VIABILITY COMPARED TO CONTROL	100 %	0.000 ug/ml	

95 materials
~16876 data points



NANoREG data availability (as of Oct 2016)

NANoREG data availability



- 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>

The screenshot displays the NANoREG search application interface. At the top, a search bar contains the query "carbon nanotube". Below the search bar, a navigation menu includes "Home", "Search", "Data collections", "Data upload", "For developers", and "Help". The main content area shows a "Hits list" with 3 items selected. The search results are displayed in a table format, with each row representing a data resource. The first three rows are highlighted with red boxes and blue arrows pointing to callouts. The first row is "JRCNM04000a (NM-400) (MWCNT 13.6 nm) carbon nanotube", the second is "JRCNM04001a (NM-401) (MWCNT 64.2 nm) carbon nanotube", and the third is "JRCNM04002a (NM-402) (MWCNT 12.7 nm) carbon nanotube". Each row includes a "more" link and a "Add to Selection" button. On the left side, there is a sidebar with "Data sources (19)" and "Nanomaterial type (92)". Under "Nanomaterial type", there are several facets or filters, including "P-CHEM (18)", "Particle size distribution (Granulometry)", "Specific surface area", "TOX (1)", "Medium (1)", "Dispersion protocol (1)", "Results", "References (19)", "Protocols (19)", and "Instruments (0)". A red box highlights the "Particle size distribution (Granulometry)" and "Specific surface area" facets. A blue box highlights the "Selected filters" callout. A blue box highlights the "List of data resources with direct links to DB" callout. A blue box highlights the "Facets or filters that permit easy refinement of search" callout.

Search query: carbon nanotube

Home Search Data collections Data upload For developers Help

NANoREG search page (demo) Search

immunotoxicity Particle size distribution (Granulometry) carbon nanotube

< 1 > displaying 1 to 19 of 19

JRCNM04000a (NM-400) (MWCNT 13.6 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

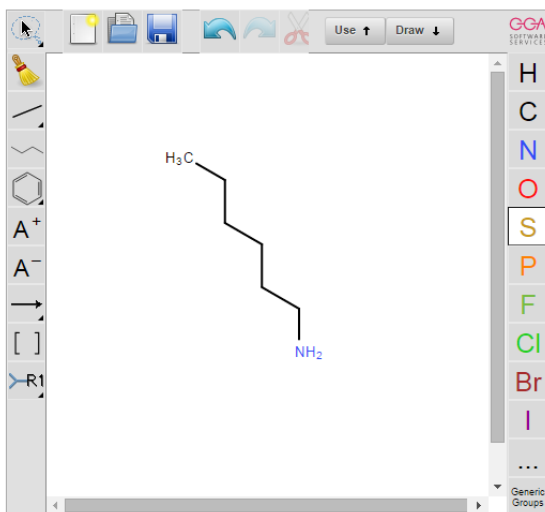
List of data resources with direct links to DB

Facets or filters that permit easy refinement of search



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 REYJPSVUYRZGE-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...	REYJPSVUYRZGE-UHFFFAOYSA-N	InChI=1S...	-	0.72

Identifiers Substances Datasets

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-ag...	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-c4f...	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

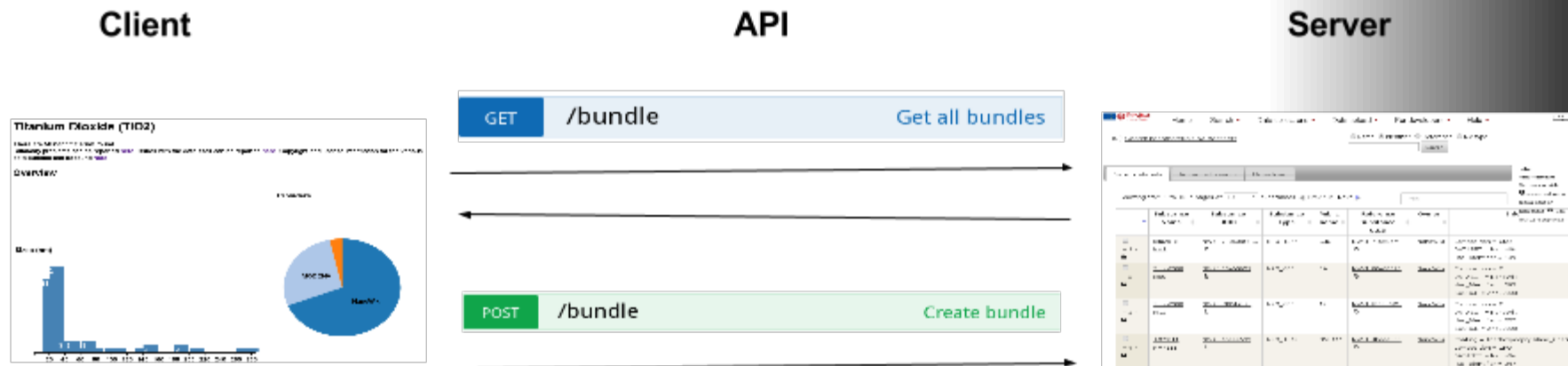
Coating

-3-			FCSV-0b...	6-Amino-1-hexanethiol	-	-	NCCCCC...	WYYXDSQPIGZPU-UHFFFAOYSA-N	InChI=1S...	-	0.62
-----	--	--	------------	-----------------------	---	---	-----------	----------------------------	-------------	---	------



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>

Integrated view of eNanoMapper database [contributors] and caNanoLab

Search

<https://data.enanomapper.net>

Hits list Selection

< 1 2 3 ... 7 8 > displaying 1 to 20 of 154

- Multi-Walled Carbon Nanotubes (MWCNT) synthetic shape mono constituent substance [12.7 nm]**
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-NJianNL2007-01 multi-walled nanotube**
P-CHEM.Particle size distribution (Granulometry) [2007]
[more](#)
[caNanoLab](#)

Reference *	Protocol	Endpoint	Result	Own	Reliability
(2011)	Transmission Electron Microscopy (TEM) DISTRIBUTION_TYPE: imaging technique	MASS MEDIAN DIAMETER	[3, 20] nm		2 (reliable with restrictions)
		MASS MEDIAN AERODYNAMIC DIAMETER	-		
		PARTICLE SIZE.D90	12.7 nm		

Showing 1 study(s) (1 to 1) [Previous](#) [Next](#)

- 4.28 Nanomaterial specific surface area (1)
- 4.29 Nanomaterial zeta potential (1)
- 4.31 Nanomaterial dustiness (1)

caNanoLab National Cancer Institute
 at the National Institutes of Health | www.cancer.gov

Navigation: HOME | PROTOCOLS | SAMPLES | PUBLICATIONS | HELP | GLOSSARY | LOGIN

Sample STANFORD-ZLiuCR2008-02

General Info	Composition	Characterization	Publication
Sample Name: STANFORD-ZLiuCR2008-02 Created Date: 1/14/11 Keywords: EX VIVO, IN VIVO, THERAPEUTIC EFFICACY, TUMOR			
Point of Contact: Yes	Primary Contact?: Hongjie Dai hdai@stanford.edu 1-650-723-4518	Contact Person: STANFORD_ChemID Department of Chemistry Stanford University Stanford CA 94305 USA	Role: Investigator

[Add to Selection](#)



- 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

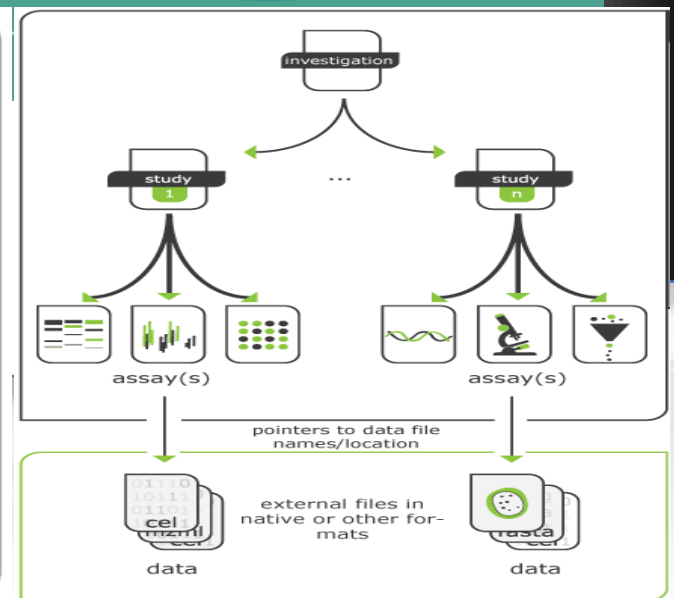
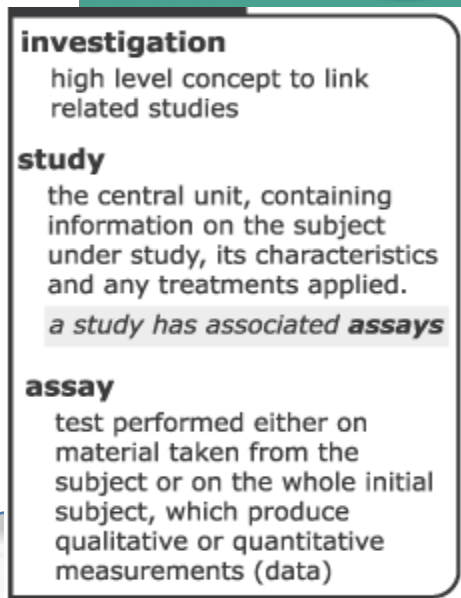
ISA-JSON version 1

(officially released 2016)

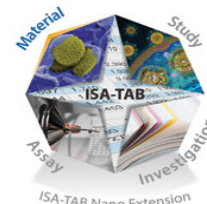
ISA-JSON Version 2

(under development)

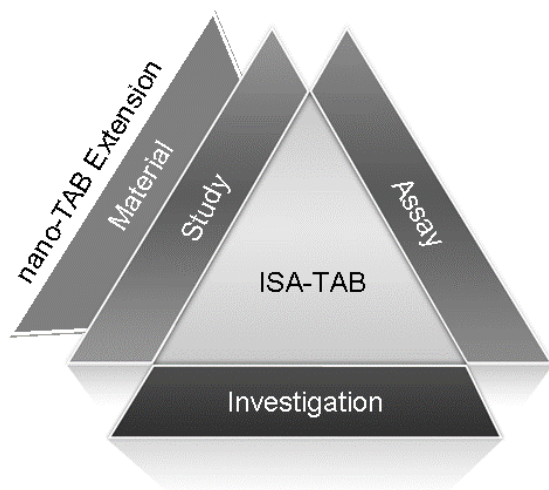
23 November 2016



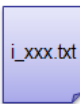
ISA-TAB-Nano



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**.

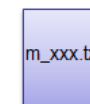
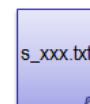


1. Describe the Investigation and Studies



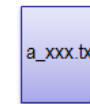
Investigation File

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)

People

10 >



ISA-JSON schemas



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

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

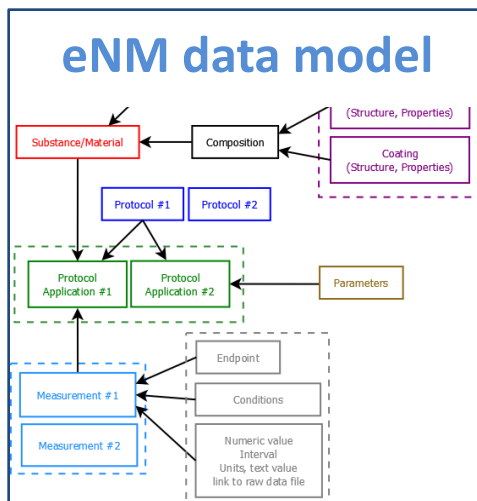
```
{
  "$schema": "http://json-schema.org/draft-04/schema",
  "title": "ISA investigation schema",
  "description": "JSON-schema representing an investigation in the ISA model",
  "type": "object",
  "properties": {
    "identifier": { "type": "string" },
    "title": { "type": "string" },
    "description": { "type": "string" },
    "submissionDate": { "type": "string", "format": "date-time" },
    "publicReleaseDate": { "type": "string", "format": "date-time" },
    "commentCreatedWithConfiguration": {
      "$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**



**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**

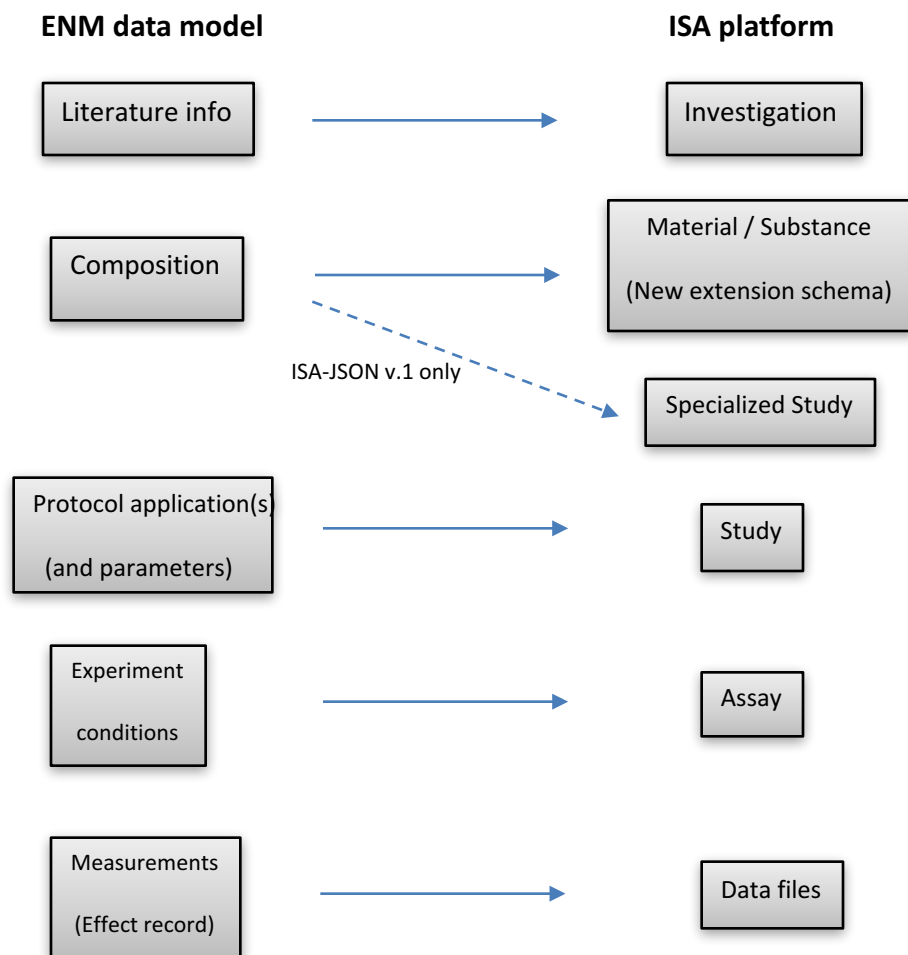
export

isa-api

ISA-TAB files



eNanoMapper – ISA mapping



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
- 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
- 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<>();
@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}
  ▶ constituent {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.

The screenshot shows the eNanoMapper web interface. The top navigation bar includes 'Home', 'Search', 'Data collections', 'Data upload', 'For developers', and 'Help'. The main content area displays a list of nanomaterial bundles. The first bundle is 'NanoWiki' (Bundle ID: 00000000-0000-0000-0000-000000000001, Version: v2). Below the table, there are 'Download as' options for XLSX, CSV, and ISA (circled in blue), and 'Export as' options for JSON, RDF, and JSON-LD (circled in green). Other bundles include 'OECD Harmonized Templates import test' and 'Protein Corona Fingerprinting Predicts the Cellular Interaction of Gold and Silver Nanoparticles'. A text box at the bottom of the screenshot contains the URL: <http://ambit.sourceforge.net/enanomapper/templates/convertor.html>

Bundle ID	Name	Ver	Description	Source	License & Maintainer
00000000-0000-0000-0000-000000000001	NanoWiki	v2	Nanomaterials, physicochemical characterisations and toxicity data, imported via NanoWiki RDF dump	NanoWiki http://dx.doi.org/10.6084/m9.figshare.1330208	http://orcid.org/0000-0001-7542-0286
00000000-0000-0000-0000-000000000002	OECD Harmonized Templates import test	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 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	www.ideaconsult.net
00000000-0000-0000-0000-000000000003	Protein Corona Fingerprinting Predicts the Cellular Interaction of Gold and Silver Nanoparticles	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 http://pubs.acs.org/doi/abs/10.1021/nn406018q	NTUA
91cad054-4de8-4dc9-a8e4-20564e7eada7	Comprehensive In Vitro Toxicity Testing of a Panel of Representative Oxide Nanomaterials: First Steps towards...	v1	Comprehensive In Vitro Toxicity Testing of a Panel of Representative Oxide Nanomaterials: First Steps towards...	http://pubs.acs.org/doi/abs/10.1021/nn406018q	www.ideaconsult.net



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>

The screenshot shows an Excel spreadsheet with two tabs: 'TEST CONDITIONS' and 'TEST RESULTS'. The 'TEST CONDITIONS' tab contains a form for entering assay details. The 'TEST RESULTS' tab displays data for four replicates, with columns for material (T1), Titanium Dioxide concentration, and Average (ng/ml).

Replicate	Titanium Dioxide	Average (ng/ml)
Replicate 1	0	1.8925
	1	19.6985
	5	18.5207
	10	18.0280
	25	18.4153
	50	19.2965
Replicate 2	0	-0.3425
	1	5.9887
	5	5.7696
	10	5.3011
	25	5.1602
	50	5.2292
Replicate 3	0	2.189
	1	2.474
	5	2.002
	10	1.786
	25	1.570
	50	1.735
Replicate 4	0	9.502
	1	8.808
	5	8.957
	10	8.306
	25	8.244
	50	9.507



Mapping the spreadsheet content into the data model

Replicate 1			Replicate 2	
T1	Titanium Dioxide	Average (ng/ml)	T1	Titanium Dioxide
	0	1.8925		
	1	19.6985		
	5	18.5207		
	10	18.0280		
	25	18.4153		
	50	19.2965		
	75	20.8867		
	100	22.6964		

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_VA...",
                      "COLUMN_POS": -1,
                      "ROW_POS": 0,
                      "UNIT": "Åµg/ml"
                    }
                  ]
                }
              ]
            }
          }
        ]
      }
    }
  }
}
```

https://apps.ideaconsult.net/enanmapper/ui/uploadsubstance1

ENM NanoMapper

Search Nanomaterials OpenTox Help

Substances > Import > Single file upload

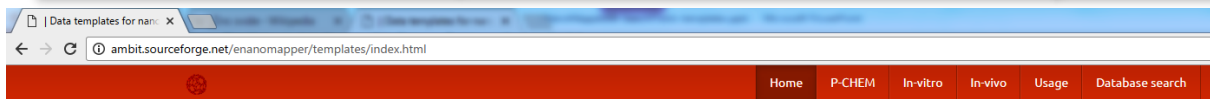
Import new substance(s)

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)



<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 https://github.com/ISA-tools/isa-api)
(R) NWrdf       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_T  
EST.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.





**Karolinska
Institutet**



misvik biology



Maastricht University

*Leading
in Learning!*

EMBL-EBI



Douglas Connect
Working communities

in silico toxicology gmbh

IOFA
consult



Questions?

THANK YOU!

