The first eNanoMapper prototype: a substance database to support a safe innovation approach



NINA JELIAZKOVA
IdeaConsult Ltd.
Sofia, Bulgaria
www.ideaconsult.net





FP7 eNanoMapper project

- Started Feb 2014, 3 years
- Develop an ontology and database unifying information about nanomaterial safety (in humans and the environment)
- Cover the full lifecycle from manufacturing to environmental decay or accumulation
- Pan-European project, 8 partners
- Ontology growth through community and re-use

Objective: Safety by Design









Maastricht University in Learning!



Douglas Connect GmbH

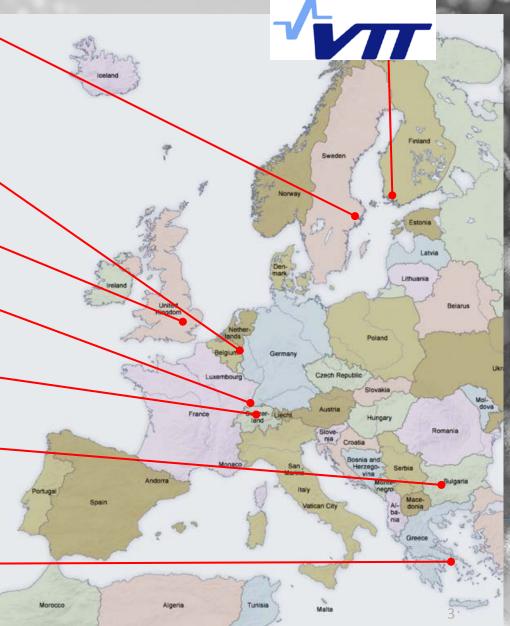






Nat. Tech. Univ. of Athens, Greece —

1/29/2015







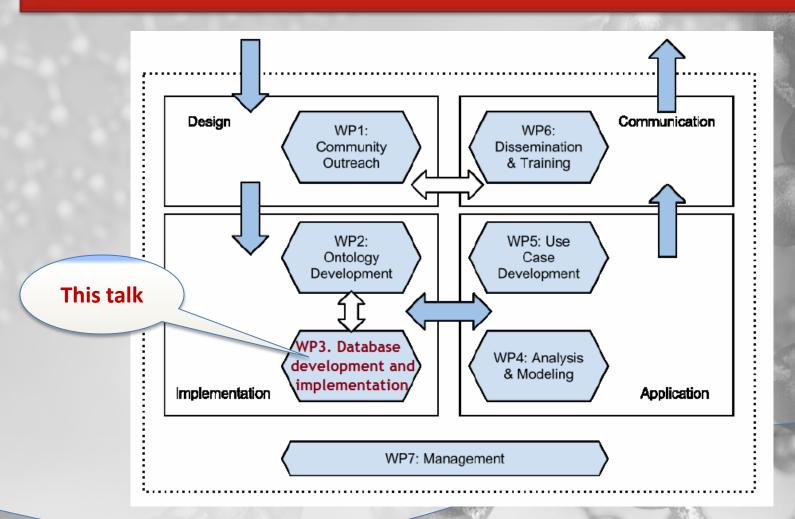
Our vision

- (re)use open ontologies
 - CHEMINF, NPO, BAO, QUDT
- Based on OpenTox API
 - EU FP7 projects: OpenTox ToxBank
 - Open Source implementations (incl. <u>ambit.sf.net</u>)
- Application Programming Interfaces (APIs)
 - Allow bridging with data analysis tools
 - Exchange formats (ISA-TAB, RDF, ...)





Work packages







eNanomapper DB review (Q1 2014)

- 104 potential data sources.
- A subset of 34 were publicly available online on the Internet.
- Most of these sources don't provide machine readable data
- Simple web pages: 18
- PDF documents: 10
- Excel tables: 3

Contributed ~30 entries to the Spring 2014 NSC Database Survey

- Database dumps : 3
- ISA-Tab-Nano format : 1
- IUCLID5 format : 1
- Semantic MediaWiki : 1
- Programmatic access through a publicly available API: 4
- Only one source makes distinction between raw and processed data and provides access to both types of data.





Nanomaterials representation

- Nanomaterials
 - Core
 - Coating(s)
 - Linkage
 - Impurities
 - Components, internal structure, etc.
- Typical assay description
 - Property value (range of values) units (Excel templates)
- More complex description:
 - Experimental graph (ISA-TAB / ISA-TAB-nano)

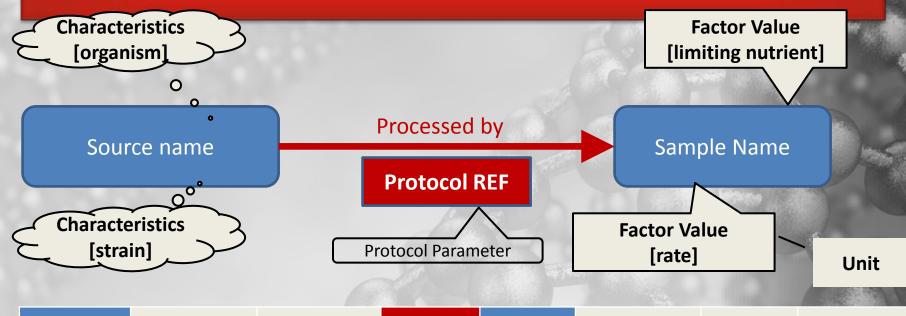
- Existing data models
 - BioAssay Ontology
 - OECD Harmonized Templates
 - CoDATA UDS
 - ISA-TAB- Nano
- Commonalities:
 - Materials sample
 - Protocols, protocol parameters
 - Experimental conditions
 - Readouts
 - Measurements,
 - Measurement groups,
 - Raw data, derived data







The Experimental graph



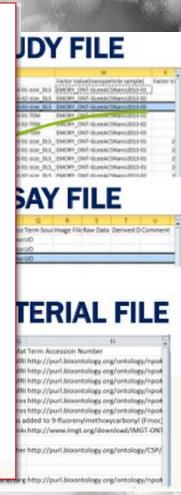
Source Name	Characteristics Characteristics [organism] [strain]		Protocol REF	Sample Name	Factor Value [limiting nutrient]	Factor Value [rate]	Unit
culture1	Saccharomyces cerevisiae	FY1679	growth protocol	C-0.07- aliquot1	carbon	0.07	l/hour
culture4	Saccharomyces cerevisiae	FY1679	growth protocol	N-0.07- aliquot1	nitrogen	Value Unit [rate]	l/hour
culture5	Saccharomyces cerevisiae	FY1679	growth protocol	N-0.1- aliquot1	nitrogen	0.1	l/hour

ISA-TAB-Nano

ISA-TAB-NANO 1.2 RELEASE NOTES

Modified the ISA-TAB-Nano 1.1 version to address user comments.

- Removed the <u>Material Linkage</u> column from the <u>Material File</u>. The <u>Material Constituent</u> column identifies the materials that are linked if the <u>Material Linkage Type</u> is specified.
- Modified the <u>Material Linkage Type</u> description to indicate that if the linkage type is an entrapment or encapsulation, the <u>Material Type</u> column can specify whether the constituent is entrapped or encapsulated.
- Rephrased Material File instructions as follows: "Materials of different chemical composition or physical characteristics should be described in separate Material files."
- Enhanced the definition of material <u>Characteristics</u> to indicate that nominal particle characteristics (or vendor supplied) should be included in the <u>Material File</u> as characteristics. Experimentally determined characteristics should be included in the <u>Assay File</u>.





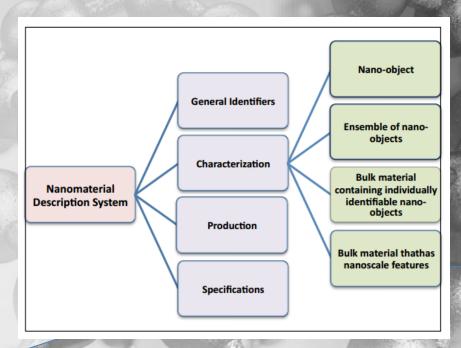


Existing data models (cont)

BioAssay Ontology

Parameter 1 measure measure has measure group' (e.g. kinase target x1...xn) Parameter 2 measure measure measure (e.g. screening concentration. c1...cm) has endpoint endpoint, e.g. endpoint endpoint endpoint percent inhibition (@ screening concentration c1) Parameter 2 (e.g. screening group 2 group 1 concentration, c) aggregated has endpoint endpoint, e.g. IC50 endpoint 1 endpoint 2 endpoint n (for kinase target x1) Legend BAO ontology class (e.g. kinase groups 1-n target x1...xn) More subclasses aggregated endpoint, e.g. BAO relation IC50 kinase endpoints 1-n profile (for kinase targets x1...xn

CODATA Uniform Description System for Materials at the Nanoscale







Experimental data

NPO

 distinguishes between endpoint of measurement and assay used to measure the endpoint, where the details of the assay could be specified

• DOI: 10.1007/s11051-013-1455-2

– "test" and "measurement" terms

CODATA UDS

requires specification of how particular property is measured.

ISA-TAB-Nano

 allows defining the qualities measured and detailed protocol conditions and instruments.

The OECD guideline

- defines the "test" or "test method" as the experimental system used to obtain the information about a substance. The term "assay" is considered a synonym. The term "Testing" is defined as applying the test method.
- The endpoints recommended for testing of nanomaterials by OECD WPMN use the terms and categories from the OECD Harmonized Templates.

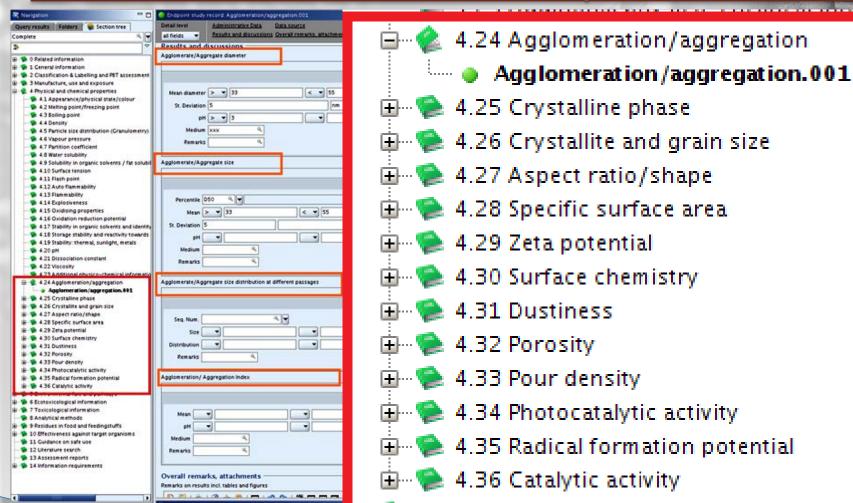
Summary

- Specify how a property is measured
- The level of details in the OECD HT, CODATA-UDS, ISA-TAB-Nano and available ontologies differ, which is due to their original focus.





Existing data models: OECD Harmonized Templates





Chemical /Toxicogenomics DB (no explicit NM support)



NM: Carbon nanotube assays >200 fullerenes; metal oxides; silver nanoparticles; colloidal gold nanoparticles, etc.

NM: Fullerenes, Metal oxides



Gene expression data

NM: carbon nanotubes, quantum dots, graphene oxide, zinc oxide, silver and gold nanoparticles.



Comparative Toxicogenomics Database Includes nanomaterial related data.



The ECHA Dissemination site. Registered chemical substances under REACH, including NM.





Q: Why no common DB approach?

- Elucidating the data model is difficult
- Making the data model universal is difficult
- Reasons:
 - Material
 - Uniqueness
 - Experimental data
 - Complexity
 - Modelling
 - Different requirements

Analogy: Chemical structures DB

- Chemical structure and properties
- Inappropriate data model.
 Instead:
 - Substances measured properties
 - Structures calculated properties.
- Substances composition
 - Constituents, impurities, additives
- Nanomaterials
 - Core, coating(s), linkage
 - Composition
 - Also impurities





What is a "Substance"

• NPO:

- a Nanomaterial (NPO_199) is an equivalent class to chemical substance (NPO_1973)
- one of (nano-object, nanoparticle, engineered nanomaterial, nanostructured material, nanoparticle formulation).
- the chemical substance itself is a subclass of a chemical entity (NPO_1972).
- DOI: 10.1007/s11051-013-1455-2 , . J. Nanoparticle Res. 2013, 15
 - Compares the definition of the terms "substance" and "material" are in ISO, REACH and general science definitions of the terms. The paper notes the OECD HT definition of "reference substances" is very similar to the definition of the term "reference material".
- REACH: http://echa.europa.eu/documents/10162/13643/nutshell_guidance_substance_en.pdf
 - "Chemical substance, a material with a definite chemical composition"
 - The definition of a substance encompasses all forms of substances and materials on the market, including nanomaterials; and may have complex composition.



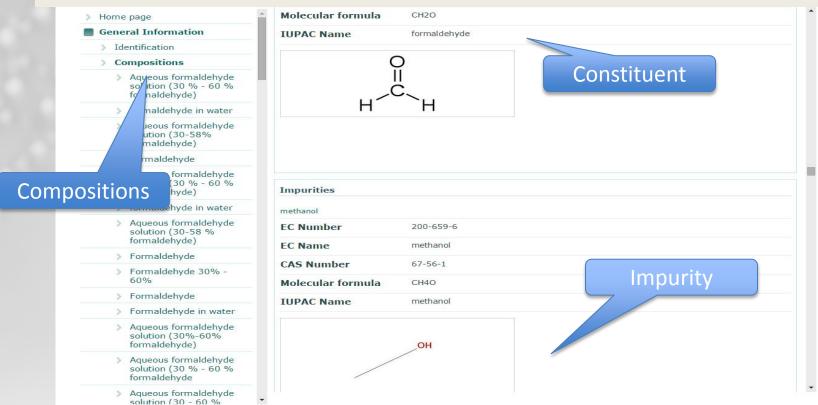


Substance (formaldehyde)

formaldehyde

Use of this information is subject to copyright laws and may require the permission of the owner of the information, as described in the ECHA Legal Notice.

European Chemical Agency Registration dossier







Substance (25155-25-3)

0733e

[1,3(or 1,4)-phenylenebis(1-methylethylidene)]bis[tert-butyl] peroxide

Use of this information is subject to copyright laws and may require the permission of the owner of the information, as described in the ECHA Legal Notice.

European Chemical Agency Registration dossier

	> Home page	Identification								
	General Information	Substance identificat	on							
	> Identification	[1,3(or 1,4)-phenylenebis(methylethylidene)]bis[tert-butyl] peroxi	ide						
	> Compositions	EC Number	246-678-3							
	Peroxide, 1,1'-[1,3 (or 1,4)-phenylenebis(1-	EC Name	[1,3(or 1,4)-phenylenebis(1-methylethylidene)]bis[tert-butyl] peroxide							
	methylethylidene)]bis[2- (1,1d-dimethylethyl)	CAS Number 25155-25-3 Reacti		Reaction mass of						
	> [1,3(or 1,4)-	Molecular formula	C20H34O4							
	phenylenebis(1- methylethylidene)]bis[tert butyl] peroxide	IUPAC Name	Reaction mass of 1,3-bis[2-(terbutylperoxy)propan-2-yl]benzene and 1,4-bis[2-(terbutylperoxy)propan-2-yl]benzene							
	> Reaction mass of 1,3-	Type of substance								
	bis[2- (terbutylperoxy)propan-	Composition	multi constituent substance							
	2-yl]benzene and 1,4- bis[2-	Origin	onstituent substance							
	(terbutylperoxy)propan- 2-yl]benzene	Trade names								
	Classification and Labelling									
1/20/2015	Manufacture, Use & Exposure		Perkadox 14-40							
1/29/2015	■ PBT assessment			7						
				,						

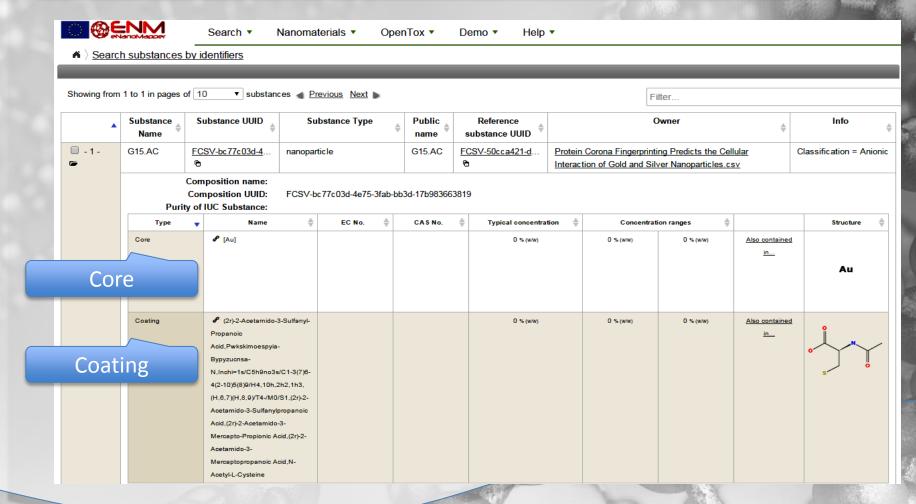
Substance (formaldehyde)

_	Substance N	lame Substance	UUID \$	Substance Type	Public name \$	Reference substance UUID	Owner \$	Info			
□ -1- ►	formaldehyde (IUC4	4 DSN 52) <u>IUC4-efdb21b</u>	<u>b-e7</u> Exist	ng Chemical		ECB5-053aa8c4-d ₾	OECD / Paris / France	IUCLID4 = DS			
	Composition name: Other: Sales Products In Aqueous Solution Composition UUID: L-cf10d721-b0e7-37cd-a233-e9a2483c4d3c Purity of IUC Substance:										
	Type	Name 🔷	EC No.	CAS No.	Typical concentration	Concentration ranges		Structure			
	Constituent	Formaldehyde	200-001-8	50-00-0	0 % (w/w)	O % (w/w)	% (w/w) Also contained in	0			
Constitu	ient										
	Impurity	P Formic Acid	200-579-1	64-18-6	ca.0.3 % (w/w)	O % (ww)	% (w/w) Also contained in	0			
Impuri	Impurity	₱ Iron.Iron(2+)	231-096-4	7439-89-6	<=0.0001 % (w/w)	O % (w/w)	% (w/w) Also contained in	Fe			
	Impurity	& Methanol	200-859-8	67-56-1	O % (w/w)	0.5 % (w/w) 2	% (w/w) Also contained in	°			
	Additive	P 0-{3-[4.6-Bis(Azanyi)-1,3,5-Triazin-2-Yi]Phenyi]-1,3,5-Triazin-2,4-Diamine,Svofitdqizskin-Uhfffaoysa-N,Inchi=1s/G12h12n10/C13-9-	225-859-0	5118-80-9	O % (w/w)	O % (www)	% (w/w) Also contained in				
	Additive (solvent)	ℰ Water	231-791-2	7732-18-5	ca.49 % (w/w)	O % (w/w)	% (w/w) Also contained in				
Additi	ive							0			





Gold nanoparticle

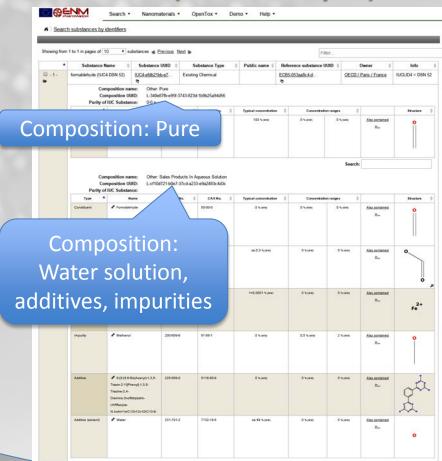




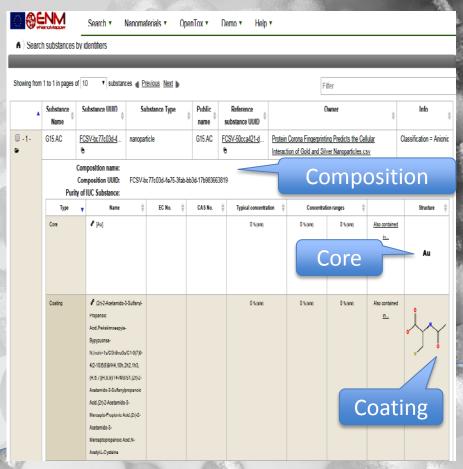


Substance, Material, Composition

Substance (formaldehyde)



Gold Nanoparticle







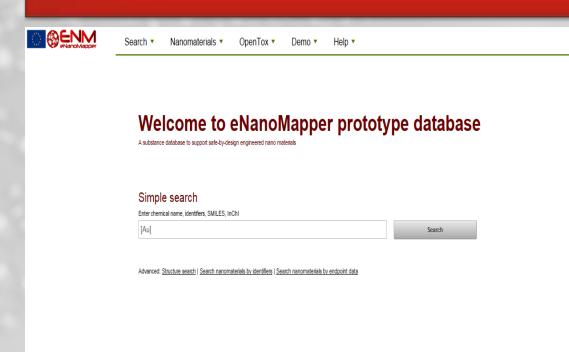
What we need to describe NM and related experiments

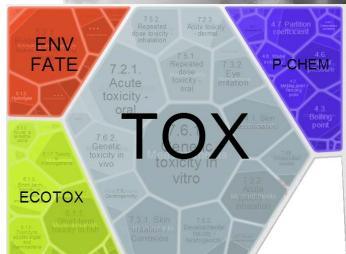
- What we need to describe a nanomaterial and related experiments
- A. An ditd o annotations, rather than an unique one. A is used to annotate samples of nanomaterials of the same type. Could also relate to an unique naming of a NM as defined by NanoDefine and FUTURENANONEEDS project.
- Csamp Leval specific to the experime
- composed of several entities (e.g. particle nucleus + coating = multiple rows ہin ISA-TAB-nano *m_* file). لغ may eventually haye different compositions
- e li ked b
- C. Attach measurements to B. A measurement is defined by applying a rtain parameters) to the sample IL. The experimental protocol P (with design might Muces as State (Confit C) factors (e.g. species, concentrations, etc.) and one of more readouts (results). The parameters, conditions are specific for given technology and endpoint measured and
- Protocol, parameters,
- Experiment readouts/resul
- rawedata-groenivecata ont?



eNanoMapper prototype database

[nina] Log out





REST API documentation: http://enanomapper.github.io/API/





Nanomaterials overview



Search ▼

Nanomaterials ▼

OpenTox ▼

Demo ▼

Help ▼

▲ Admin > Statistics > Substances per owner

Showing 1 to 4 of 4 entries

Substance contributor Name	Substances	Table view	
Ideaconsult Ltd. / Sofia / Bulgaria	Substances [1]	Substances and data Chemical structures	IUC5-354430
NanoWiki	Substances [330]	Substances and data Chemical structures	NWKI-9F4E8
OECD / Paris / France	Substances [1]	Substances and data Chemical structures	IUC4-44BF02
Protein Corona Fingerprinting Predicts the Cellular Interaction of Gold and Silver Nanoparticles.csv Protein Corona	Substances [121]	Substances and data Chemical structures	FCSV-31961

Show

10 ▼

entries



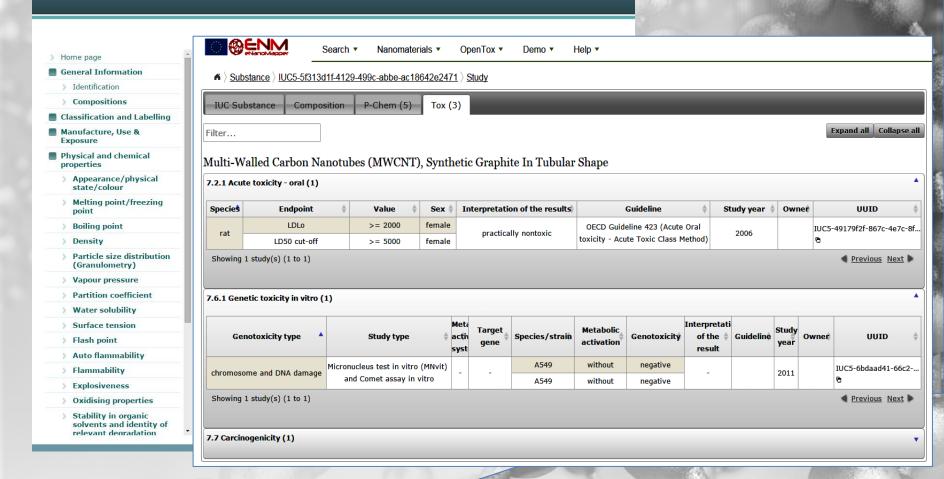
dataset



Multi-Walled Carbon Nanotubes

Multi-Walled Carbon Nanotubes (MWCNT), synthetic graphite in tubular shape

Use of this information is subject to copyright laws and may require the permission of the owner of the information, as described in the ECHA Legal Notice.







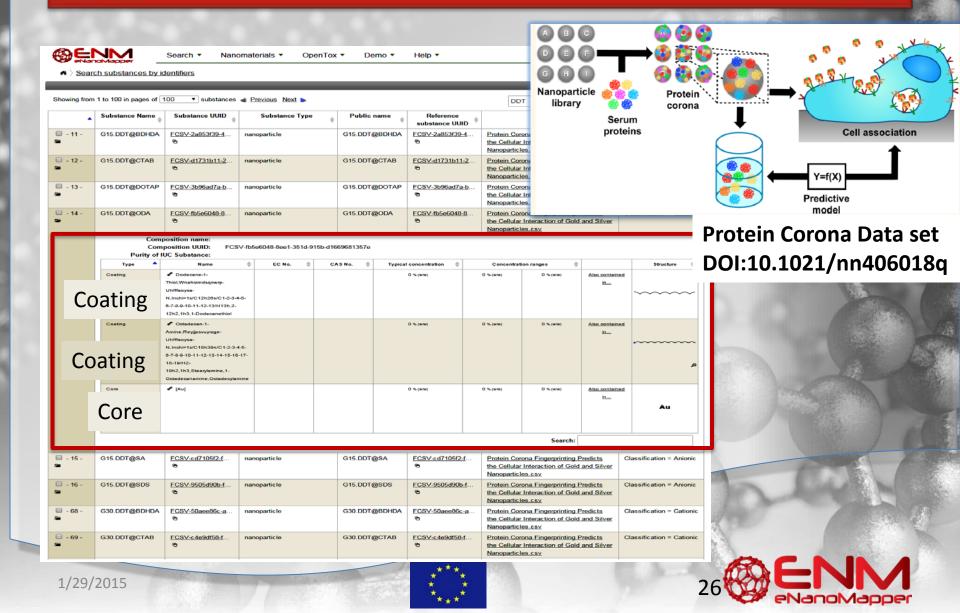
MWCNT (Phys-chem)

•		Searc	ch ▼ N	Nanomateria	ls ▼	OpenTox	▼ Demo	▼ Help	, ▼								
☆ 〉 <u>Sub</u>	stance > IU	C5-5f313d1f-4	129-499c-	abbe-ac186	42e247	L > Study											
IUC Su	bstance	Composition	P-C	hem (5)	Tox (3	3)											
lter																Expand all Co	llapse a
ulti-W	alled Ca	rbon Nanot	ubes (M	IWCNT),	Synth	etic Gra	phite In Tu	bular Sł	ape								
1 Appea	rance (1)																
Ph	ysical state		Re	mark	\$		Substance ty	/pe	\$	•	Guideline	\$		Owner	\$	UUID	
	solid		nano	material			inorganic									IUC5-9db5f020-e0fa	-47c8-8
howing	1 study(s) (1 to 1)														◀ Previous	Next >
5 Partic	le size disti	ribution (Granu	ulometry)	(1)													
Test	Dist Passa																
ateria	type num.	End		int \$			Value		dev. \$	Reference		Guideline \$		Owne	UUID		
	ima -	MAS	S MEDIAN	DIAMETER		1	[3, 20)		-								
-	ging -	MASS MEDIAN AERODYNAMIC DIAMETER PARTICLE SIZE.D90					-	2011		Transmission Electron Microscopy (TEM)			IUC5-4bca14fc-3e43-435e				
	hniq ue			= 12.7			-		Therese		гозсору (1	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,					
howing	1 study(s) (1 to 1)														◆ Previous	Next >
28 Nand	omaterial s _i	pecific surface	area (1)														
29 Nano	omaterial ze	eta potential (1)														
Type of nethod		Endpo	oint \$	Resu	lt \$	Remarks	Std. dev. ♦		н	\$	Medium	⊕ R	eference -	Guidelin	ê Ownê	UUID	,
		ZETA POT	ENTIAL	= 41.	4	-	= 2.9	=	1.9		-						
	_	ZETA POTENTIAL = 39.3		3	-	= 1.5	=	2.99		-	2010				IUC5-708684e4-4e3b-4	b-49cd-	
		ZETA POT		= 21.	4	-	= 6.5	= -	1.26		-	2010				6	
		ZETA POT	ZETA POTENTIAL		5	-	= 1.7 = 5.04		5.04	-							
howing	1 study(s) (1 to 1)														◆ Previous	Next >
31 Nano	omaterial d	ustiness (1)															

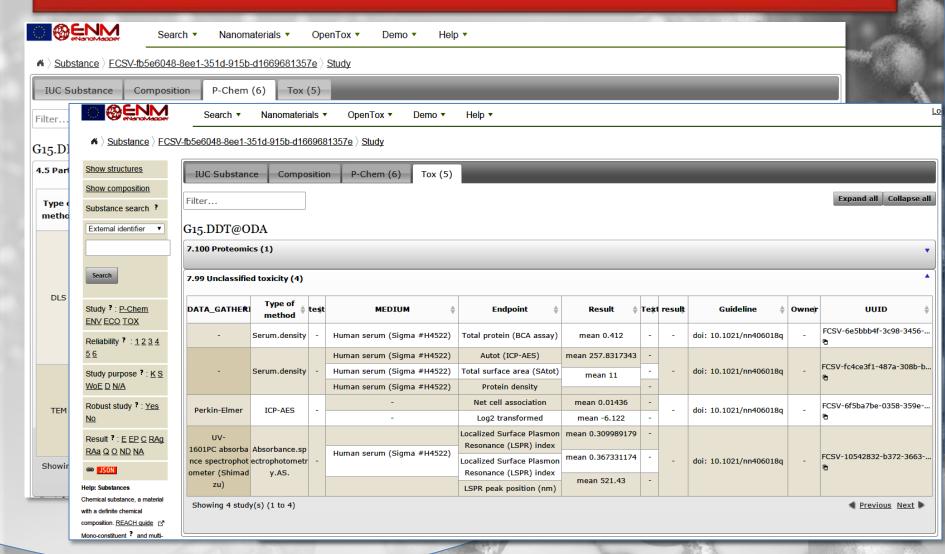




Prototype database (Protein Corona)



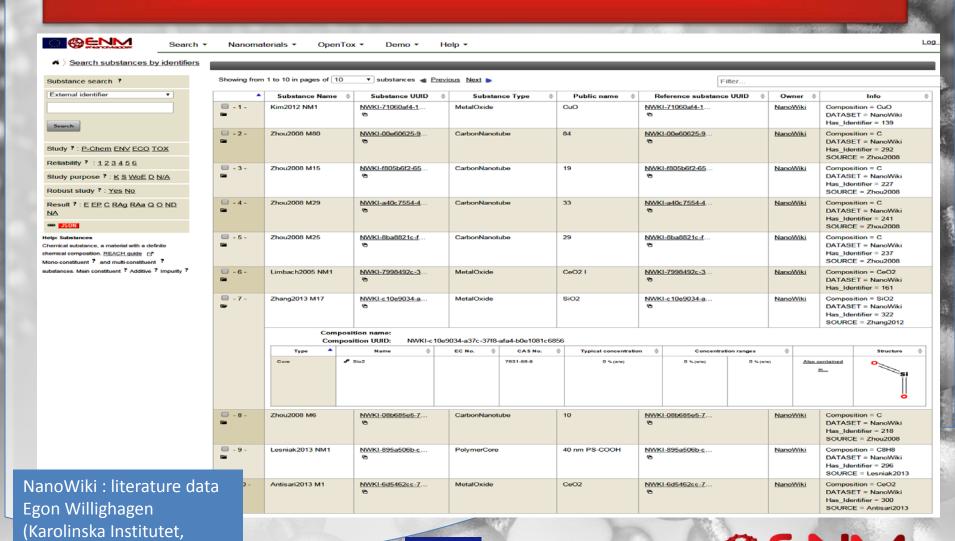
Prototype database (Protein Corona)





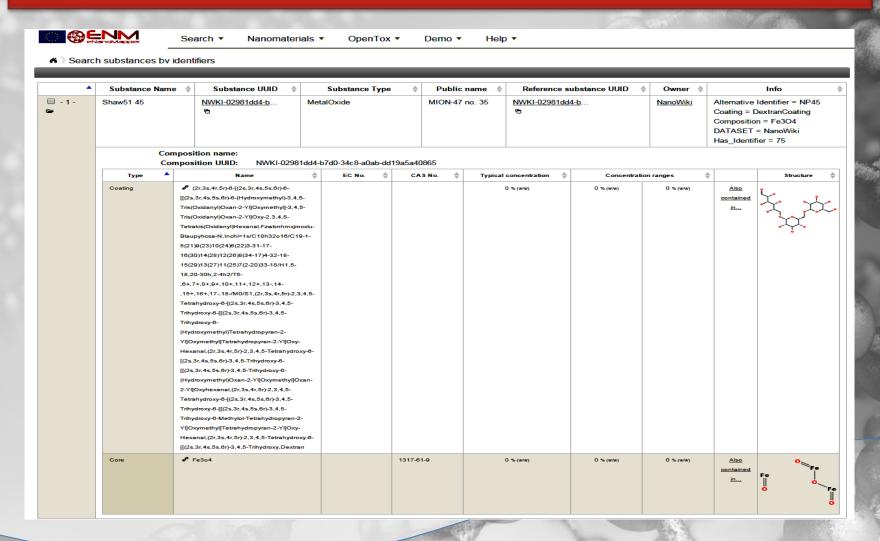


Prototype database (NanoWiki)



Maastricht University)

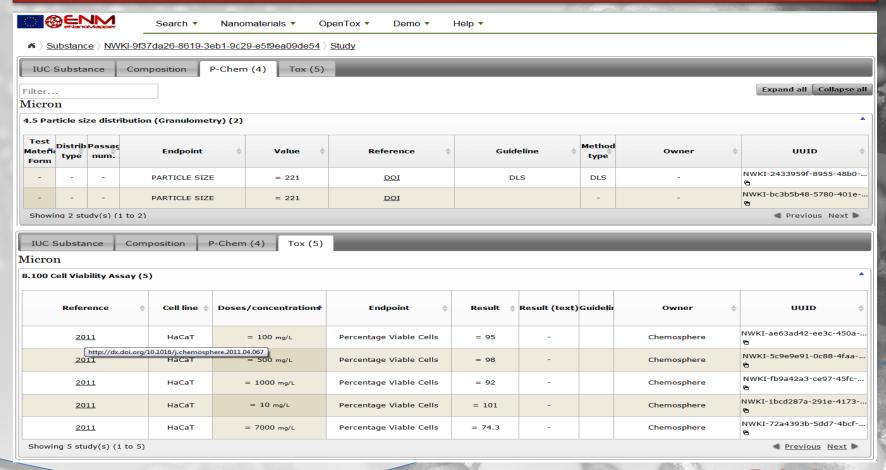
NM components (NanoWiki)







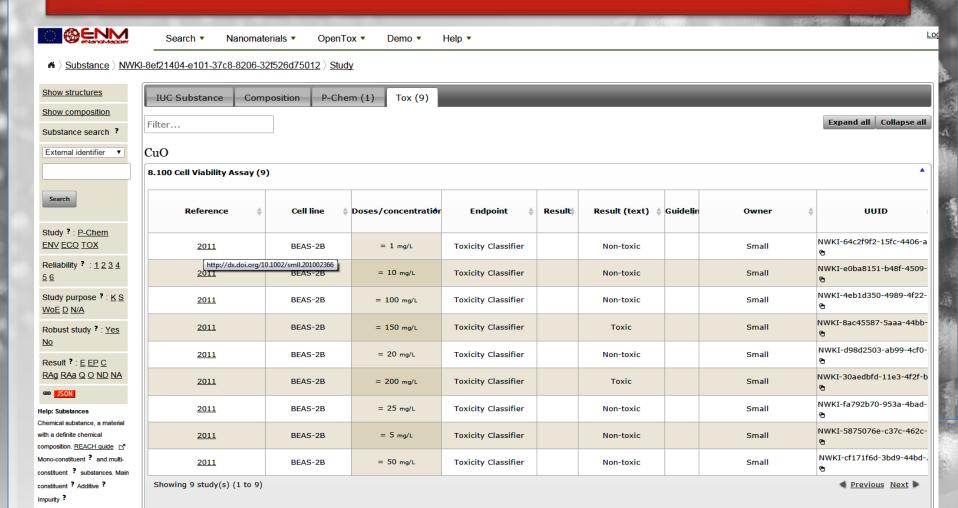
Prototype database: Phys chem & Tox (NanoWiki)







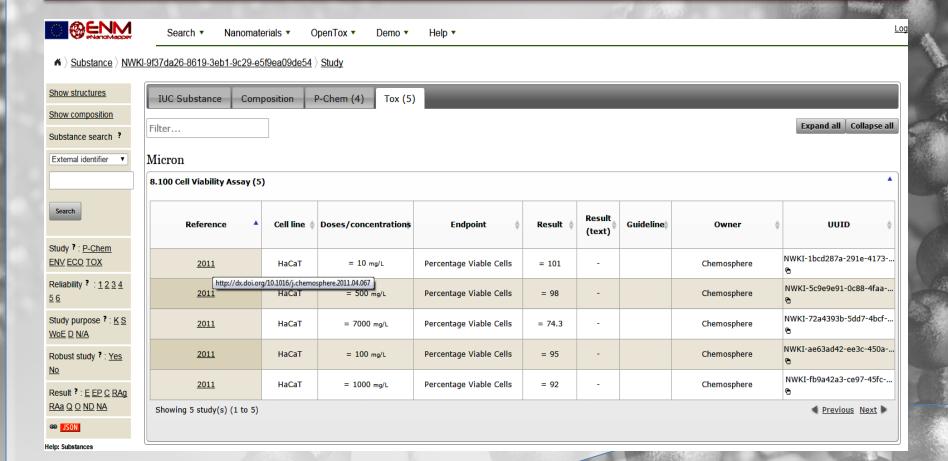
NM toxicity (classifier)







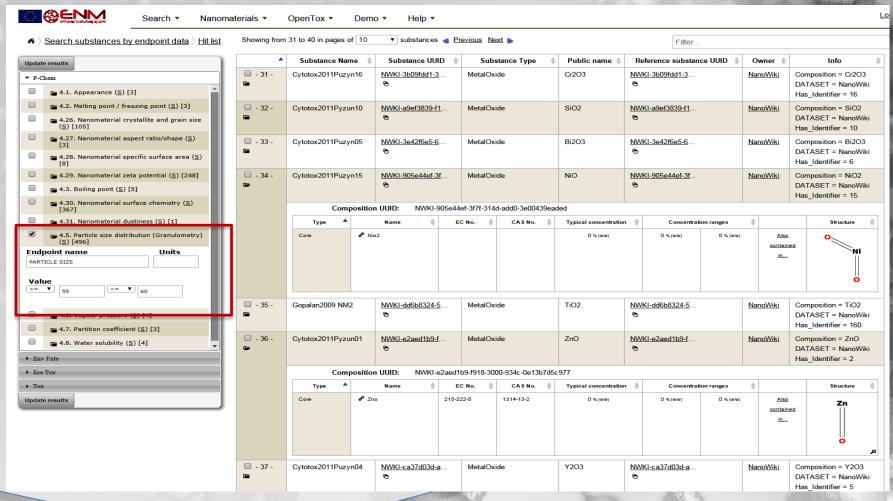
NM toxicity (numeric)







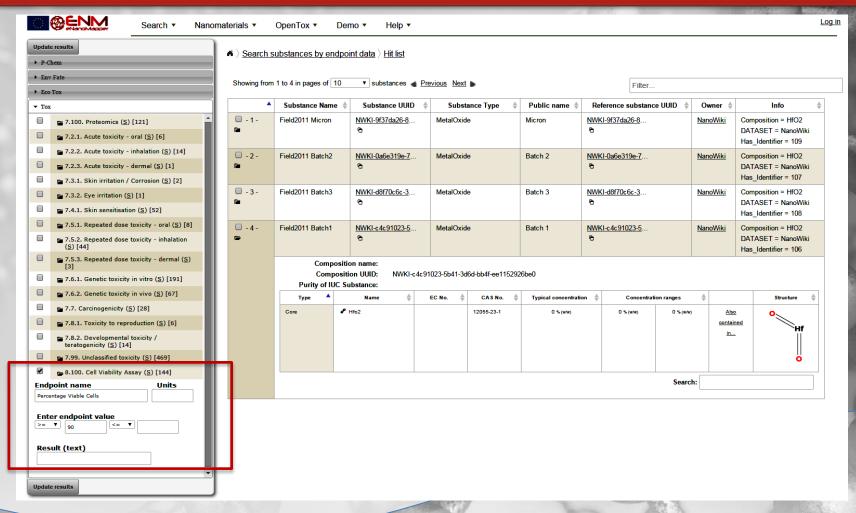
Search by particle size







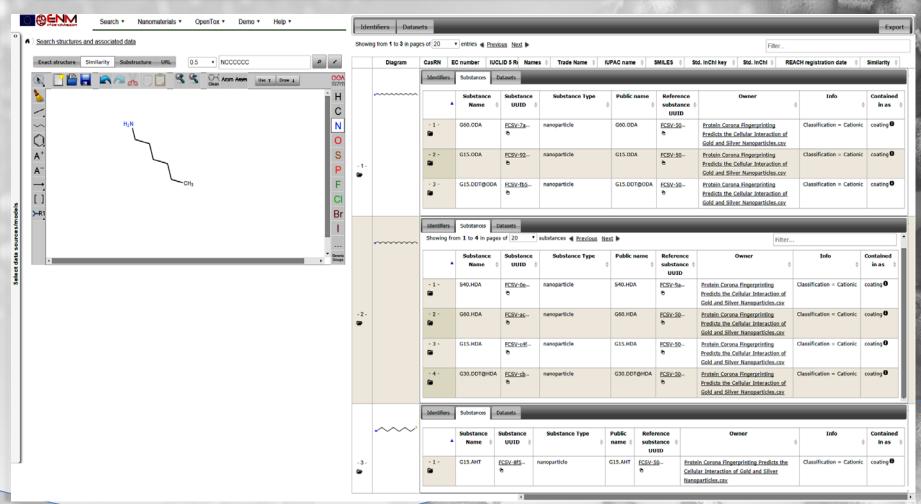
Endpoint search (tox & pchem)







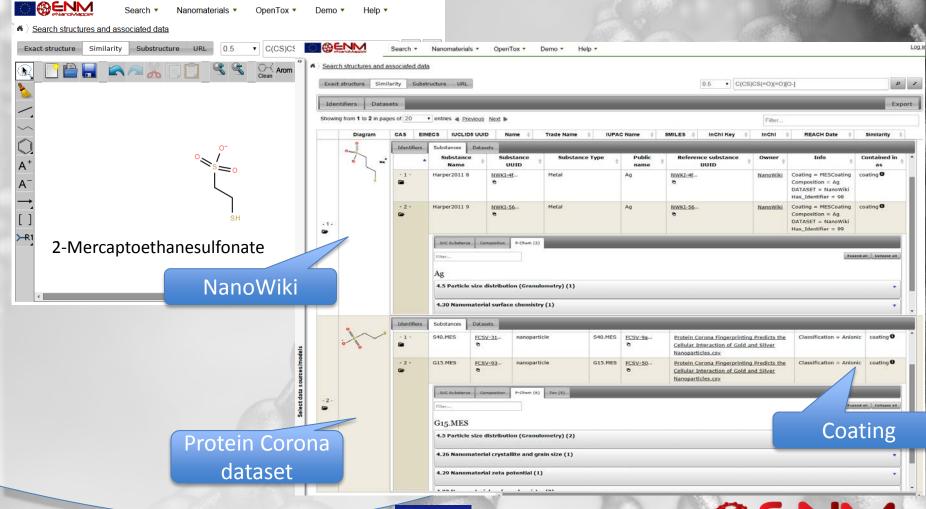
Chemical similarity search







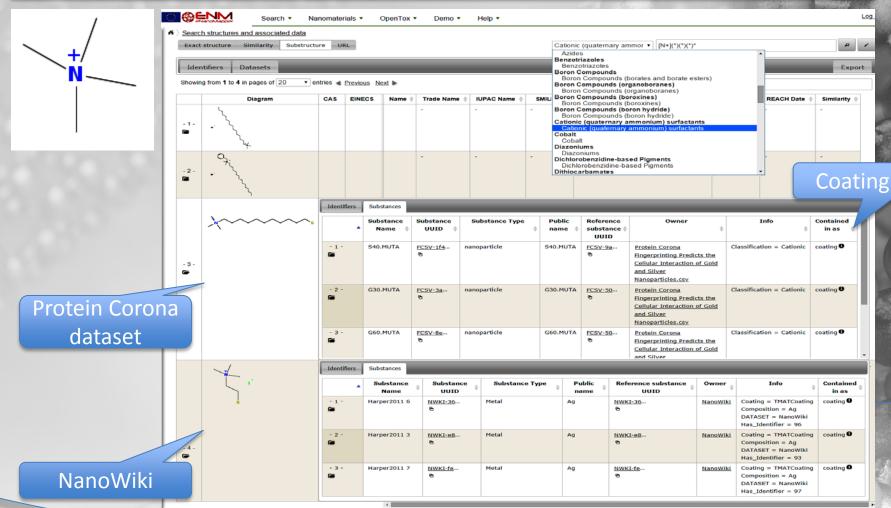
Chemical similarity search (MES coating)







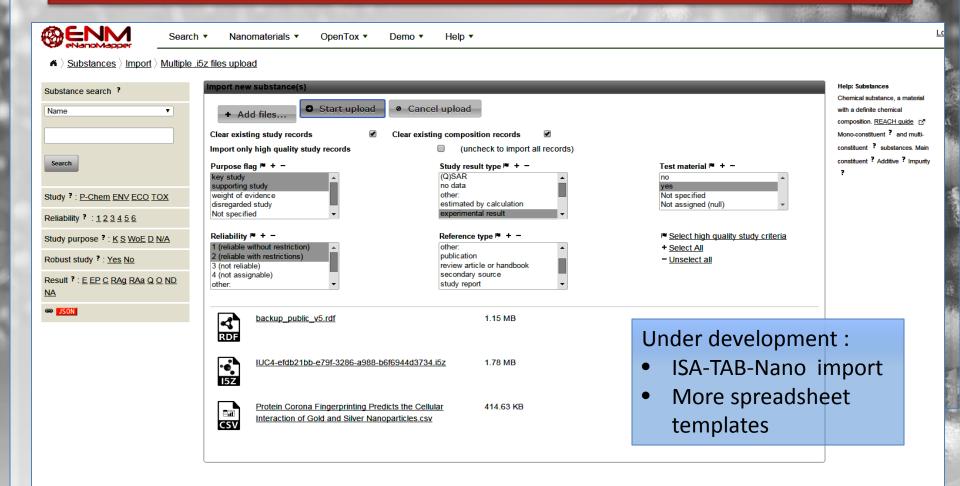
Chemical substructure search







Data upload



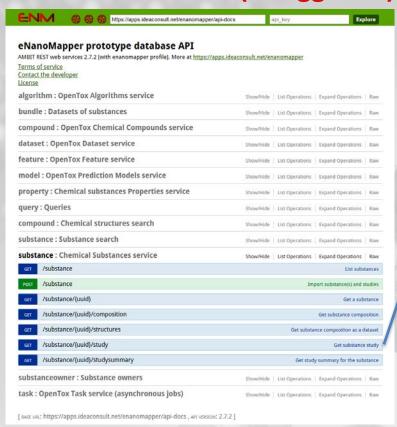
Substance import options: Multiple _i5z files upload | Single _i5z file upload | Retrieve substance(s) from IUCLID5 server





REST Application Programming Interface

API documentation (Swagger-UI)



http://enanomapper.github.io/API/

Interactive API queries



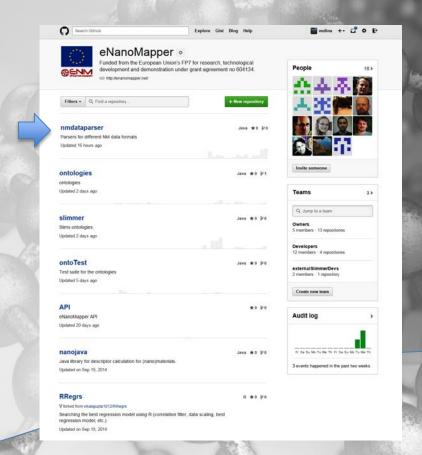


Data formats

Import formats

- Available
 - Custom CSV
 - Custom RDF (NanoWiki)
 - IUCLID5 *.i5z
- Under development :
 - ISA-TAB-Nano parser
 - Configurable parser for spreadsheet data templates, provided by NSC projects

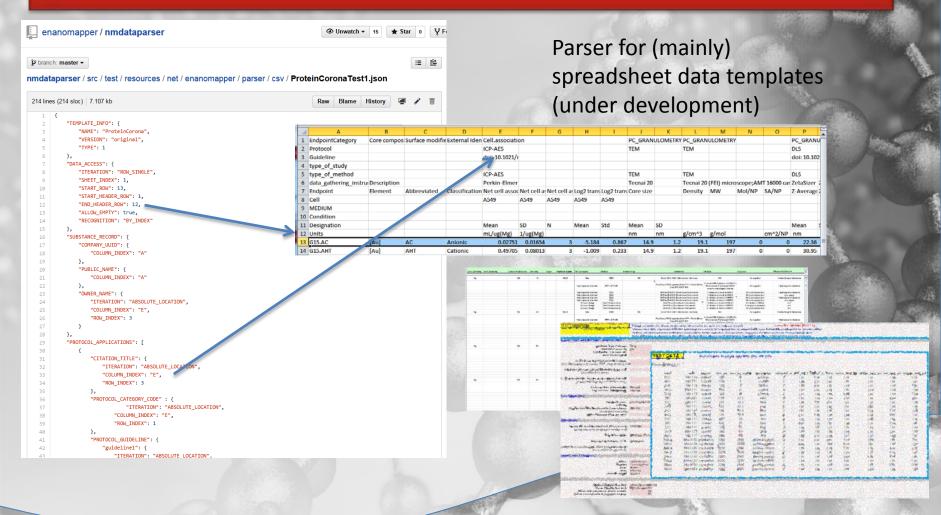
eNanoMapper@gitHub







Data formats





Questions? THANK YOU!