ENM TUTORIALS

eNanoMapper database : data access and upload

RELEASE DATE:	6.2.2017
USE:	Data Warehous; Data Analysis
VERSION:	V.4.0
MAIN AUTHOR:	Nina Jeliazkova, Egon Willighagen
PARTNER:	IDEA, UM
CONTACT DETAILS:	partners@lists.enanomapper.net
AUTHORS:	Nina Jeliazkova, Egon Willighagen
LICENCE:	CC-BY 4.0

TABLE OF CONTENTS

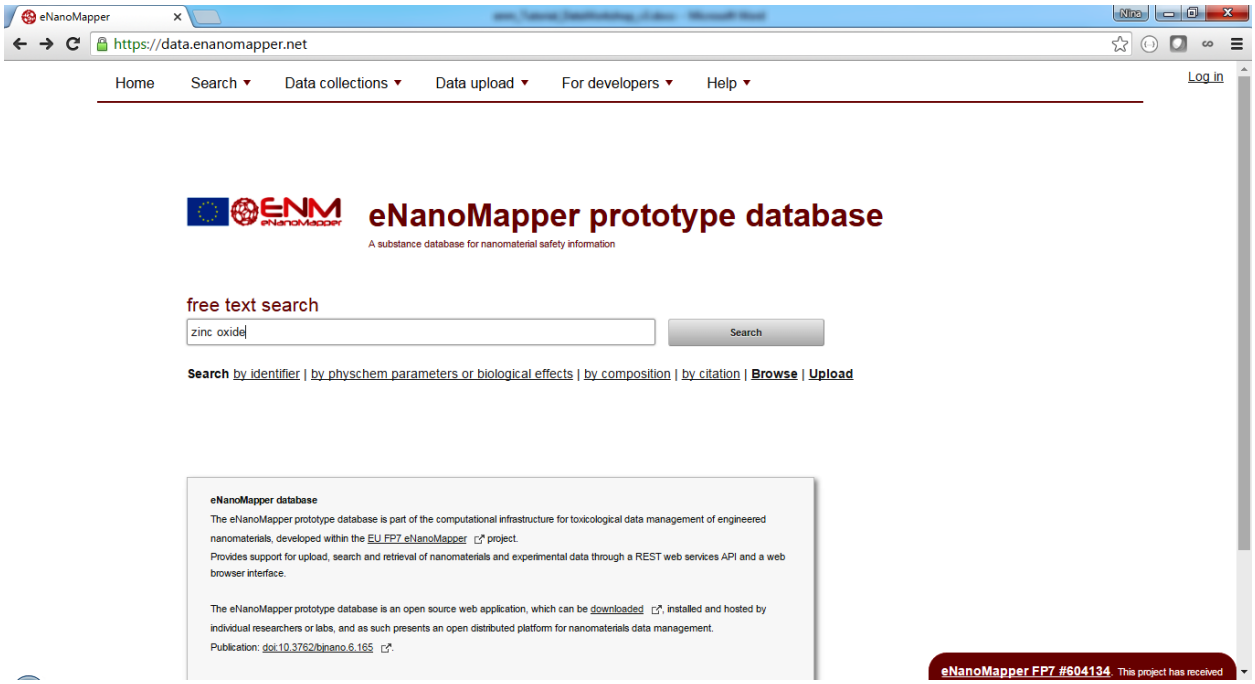
1. INTRODUCTION

The goal of this workshop is to make the participants familiar with the eNanoMapper solutions for data management and data access. We will demonstrate how the http://data.enanomapper.net/ integrates various data sets, how you can search for materials, how you can upload data, and how we can use the application programming interface (API) to access data. This document provides information how to run the exercises. For detailed description of the eNanoMapper data solutions, please consider the publication doi:10.3762/bjnano.6.165 and resources at http://enanomapper.net/.

2. SEARCH

The task outlined in this chapter is to search for zinc oxide nanomaterials.

2.1. GO TO HTTP://DATA.ENANOMAPPER.NET



2.2. ENTER "ZINC OXIDE" AND CLICK ON "SEARCH" BUTTON

This launches the search application you will explore. The page shown in Fig. 1 appears.

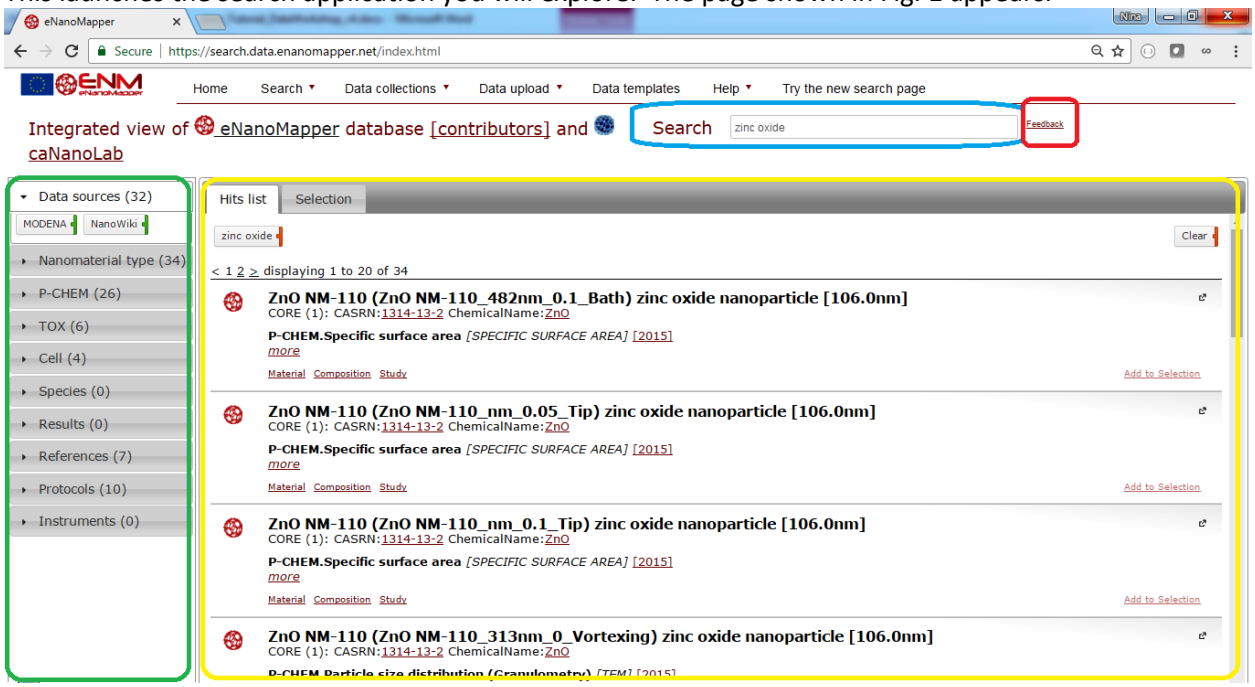


Figure 1 Zinc oxide search.

- There is a search box (highlighted in blue, top) a summary panel (left, green) and the results are shown at the main panel (yellow).

2.3. TYPE IN THE SEARCH BOX "ZINC OXIDE" AND PRESS ENTER.

The result page as in Fig.1 will appear. There are several summaries at the left panel as shown on Fig 1.

- Data sources
- Nanomaterial type
- P-Chem
- Tox
- Cell
- Species
- Results
- References
- Protocols
- Instruments

Every panel is expandable and shows the types of elements found for the particular query, "zinc oxide" in this case. For example, there are two data sources shown, "MODENA" and "NanoWiki", because these two data sources contain entries for zinc oxide particles.

2.4. CLICK ON NANOMATERIAL TYPE AT THE LEFT

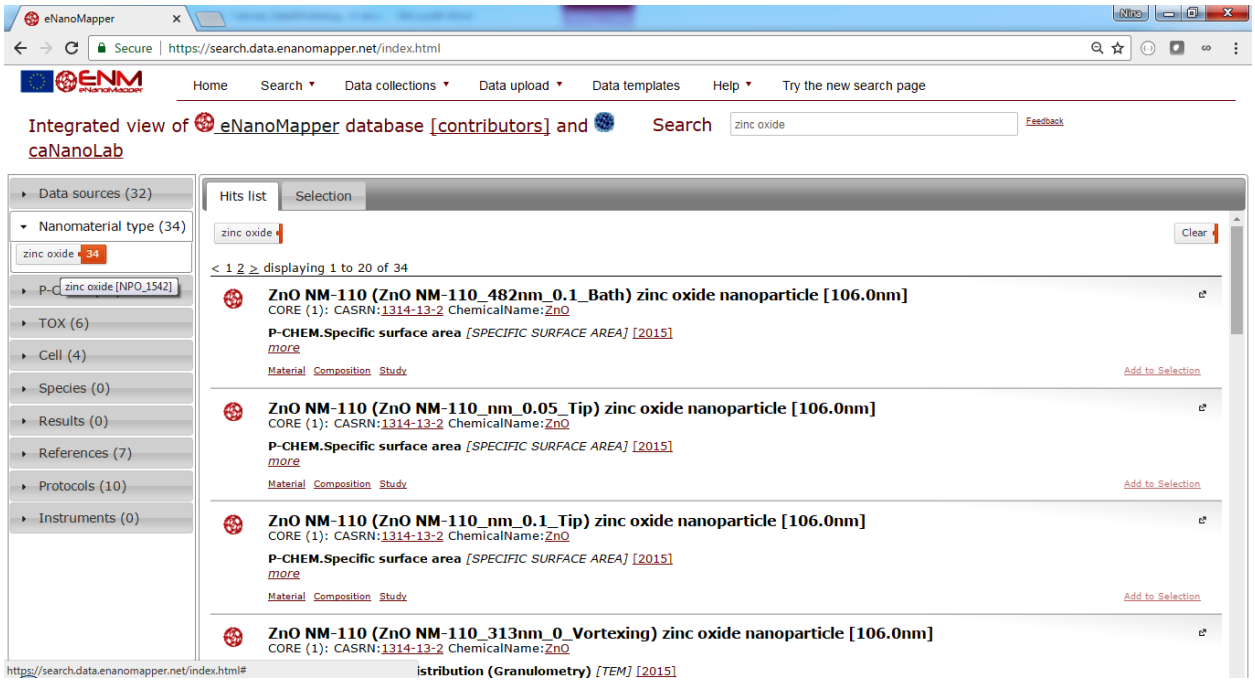


Figure 2 The nanomaterial type panel shows zinc oxide only, because this is what the query is about.

2.5. CLICK ON P-CHEM PANEL AT THE LEFT.

P-CHEM stands for physico-chemical characterisation and shows a summary of the type of experiments (the tags marked with green line at the right) and endpoints (the tags marked with blue line at the right).

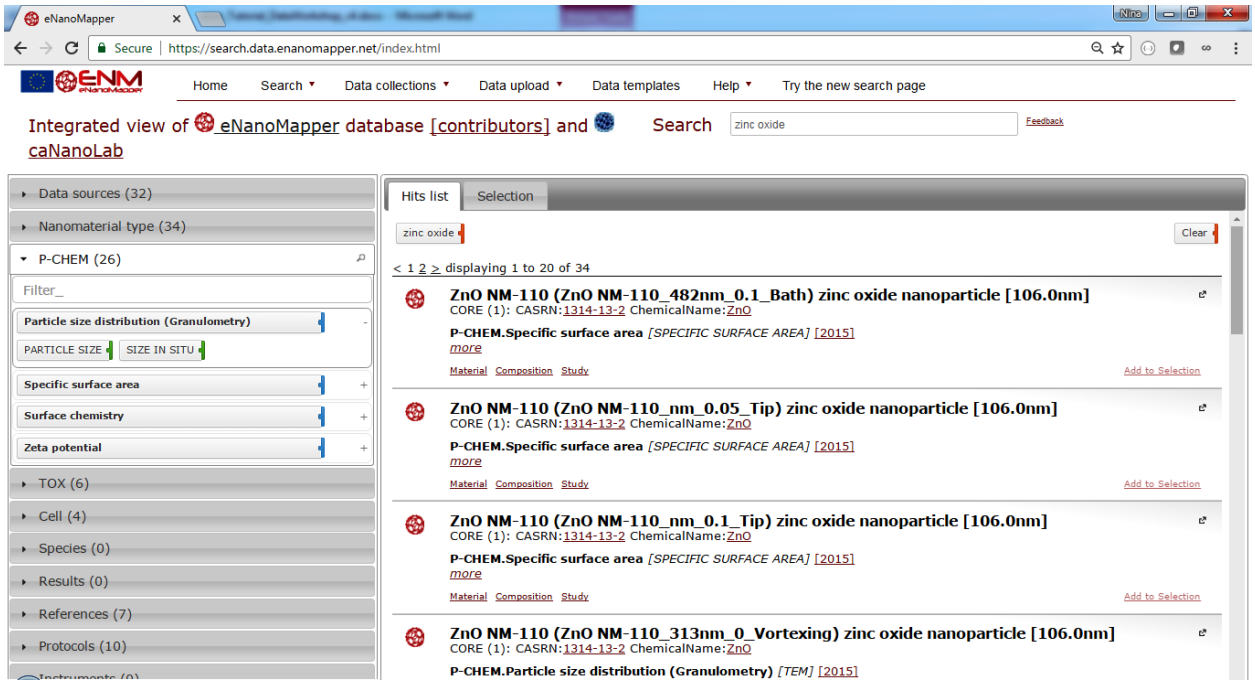


Figure 3 The P-CHEM (physico-chemical characterisation) panel shows the type and the number of entries available for different physicochemical measurements.

Mouse hovering on each tag reveals more information, as number of entries (the colored part of the tag) or ranges of the available measurement (tooltip on the tags marked blue), see Fig 4.

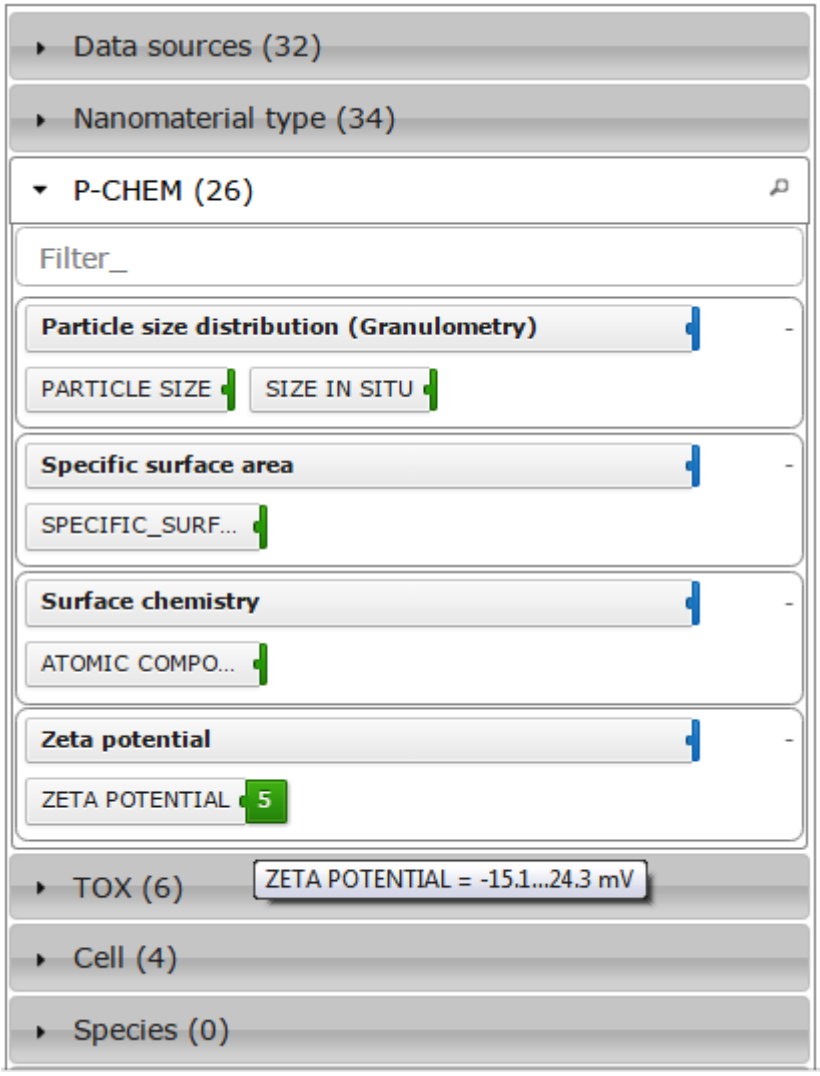


Figure 4 Summary details, e.g. shows the range of the Zeta potential measurements (-15.1, 24.3) mV

2.6. CLICK ON TOX PANEL AT THE LEFT.

TOX stands for toxicity assays, and shows a summary of the type of the experiments (the tags marked with green line at the right) and endpoints (the tags marked with blue line at the right).

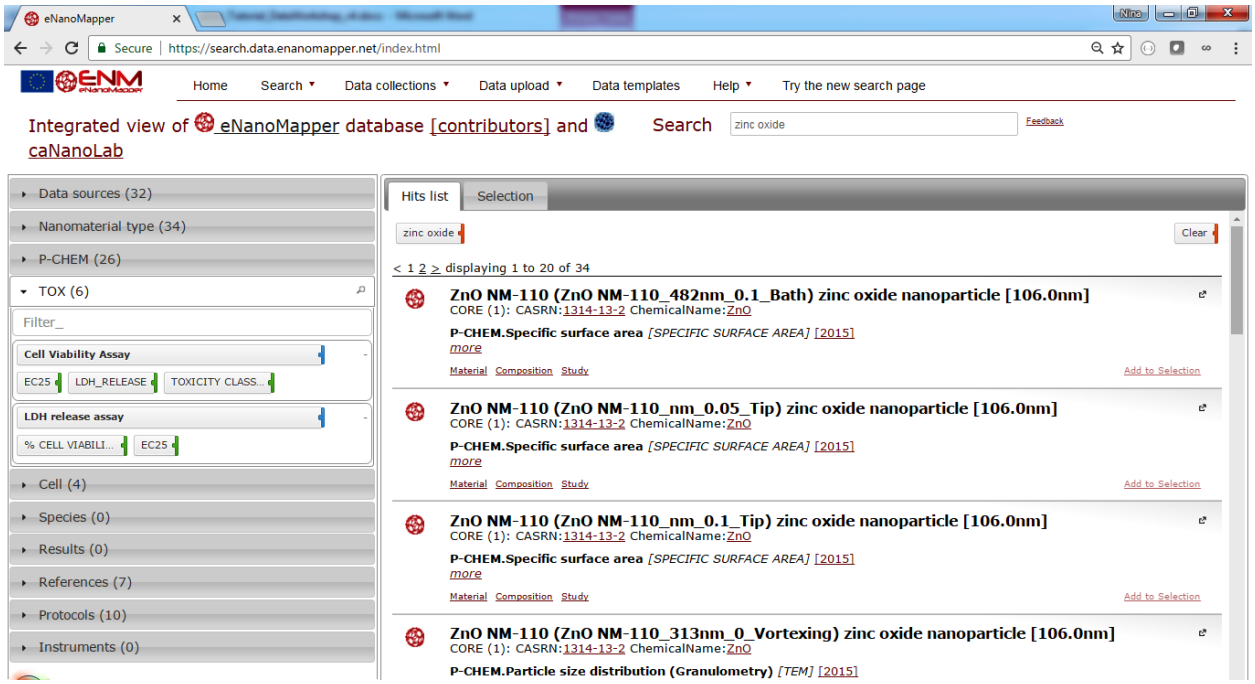


Figure 5 The TOX (toxicity assays) panel shows the type and the number of entries available for different biological assays. Hover with mouse on each tag in order to see more details.

2.7. CLICK ON *CELL* PANEL ON THE LEFT

The expanded panel will show the cell lines used in the biological assays involved with zinc oxide (recall we are still exploring the "zinc oxide" search results, and the current filter is indicated by tag 'zinc oxide' at the top of the hit lists).

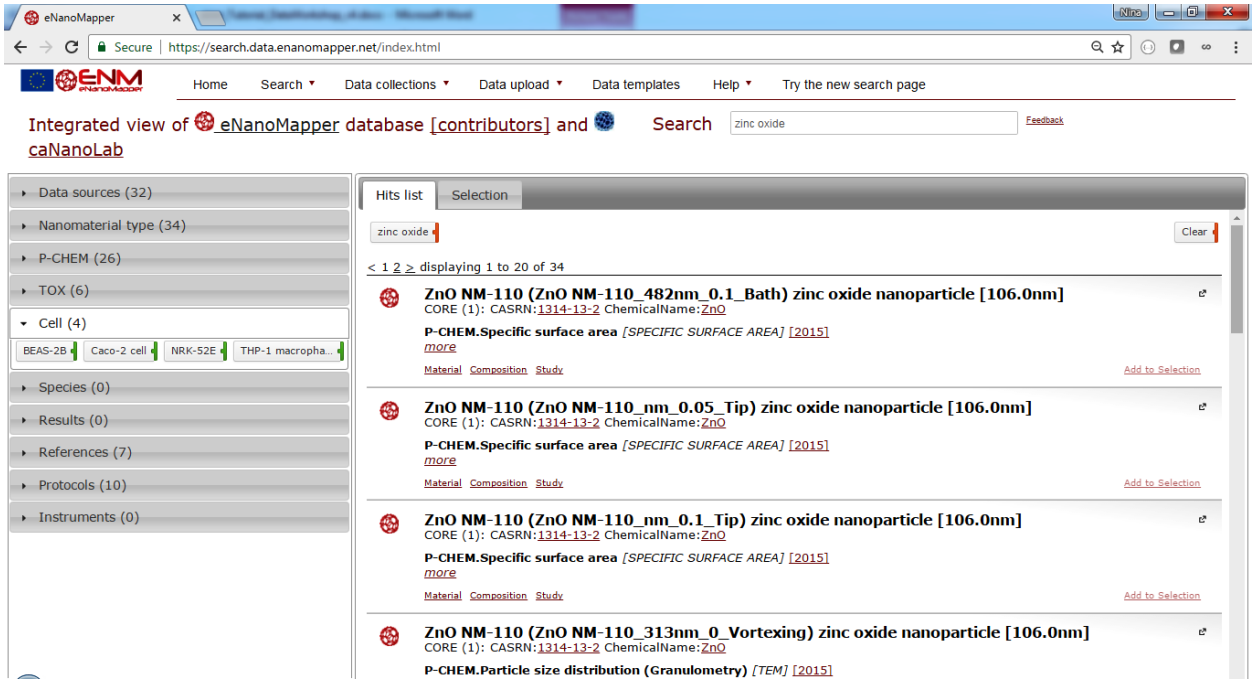


Figure 6 The cell lines used in the experiments, involving zinc oxide nanoparticles.

2.8. CLICK ON THP-1 MACROPHAGES BUTTON

The *current filter* will be updated with a second entry, restricting the query for *zinc oxide* + *THP-1 macrophage*. This is an illustration one can restrict the query by clicking any combination of the tags from the summary panels. The content of the result and summary panel will adapt to reflect the query results.

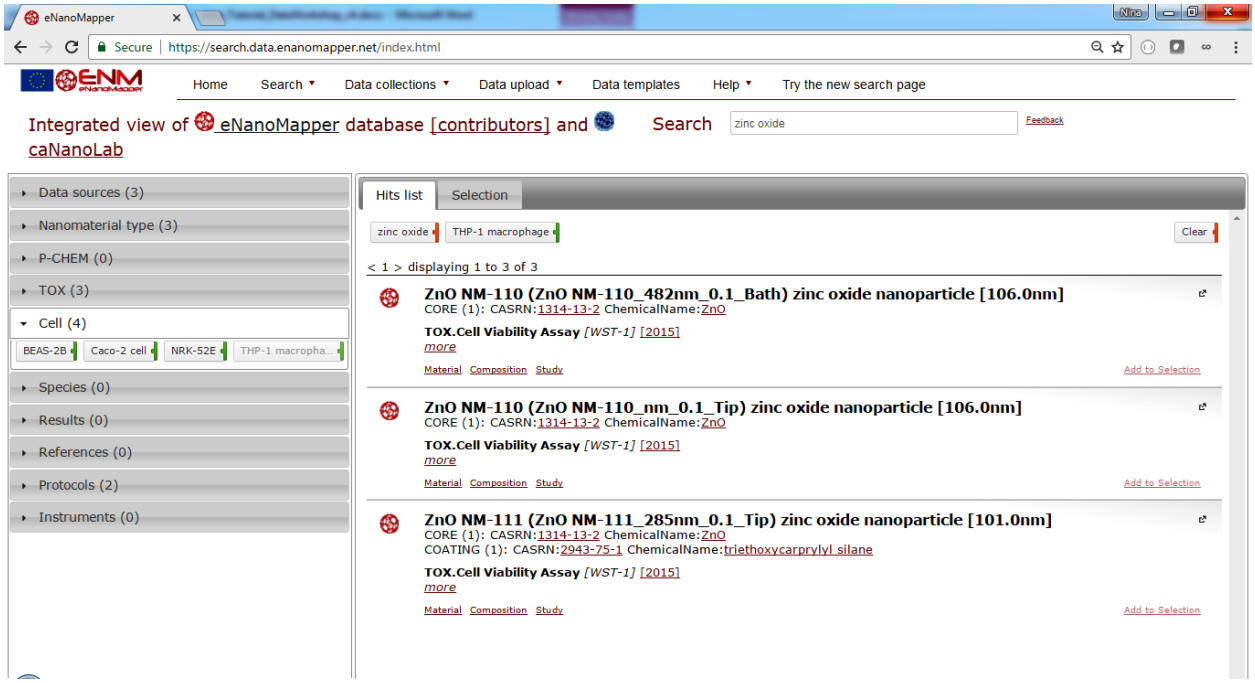


Figure 7 The cell lines used in the experiments, involving zinc oxide nanoparticles.

The content of the *current filter (at the top of the hit list)* can be directly modified by clicking on the tags or using the button *clear,* which will remove the selection and the result and summary panel content will be updated.

2.9. REMOVING FILTERS

Clicking on the *zinc oxide* tag at the top of the hit list will remove the *zinc oxide* from the query, retaining only the *Cell line : THP 1 macrophage* criteria. The results panel and the summaries will be adapted accordingly.

As the results list is now updated, it contains particles other than *zinc oxide*. Click on the *Nanomaterial type* to explore what particles are included in the search results.

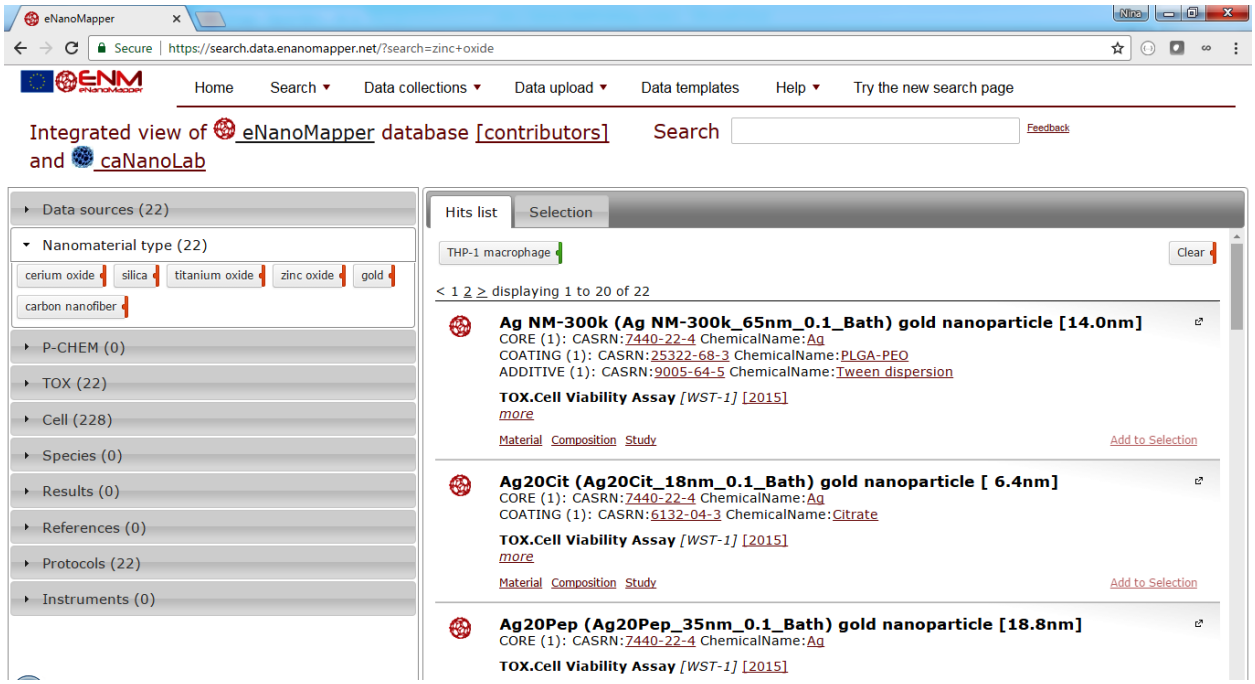


Figure 8 Click on the nanomaterial type panel on the left to see the summary of nanomaterial types found in the current query.

2.10. PROVIDE FEEDBACK

You can click on the *Feedback link* (top right, marked with red), or the links marked with blue to explore the full study records of nanomaterials (Fig 10). Feedback can be issues you found, comments, and requests for further information.

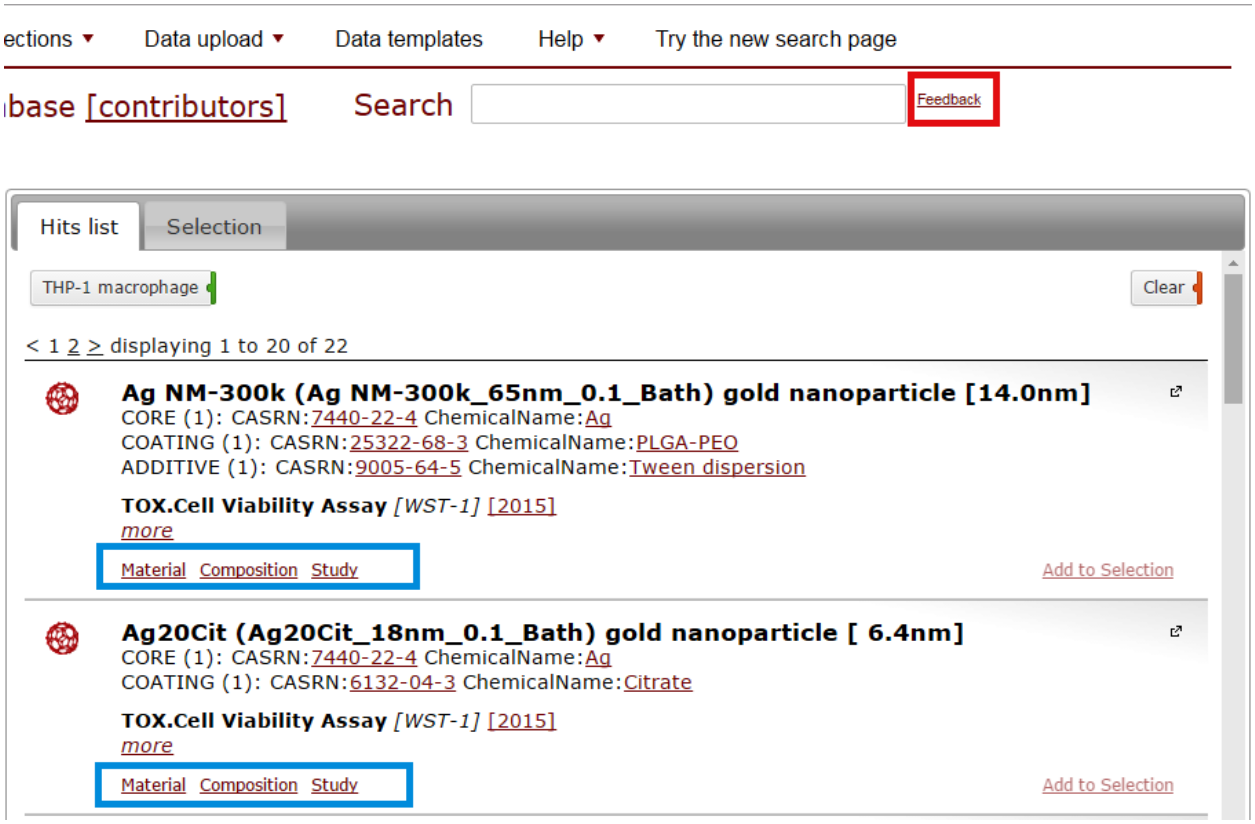


Figure 9 Links marked blue will lead to a database page exploring the full record of the nanoparticle characterisation. The link marked red is a feedback form with three questions

2.11. START A NEW SEARCH WITH THE NEW VERSION OF THE SEARCH

We have recently released a new updated version of the search page, with several new and improved features. To access the new version, use the "Try the new search page" at the top menu.

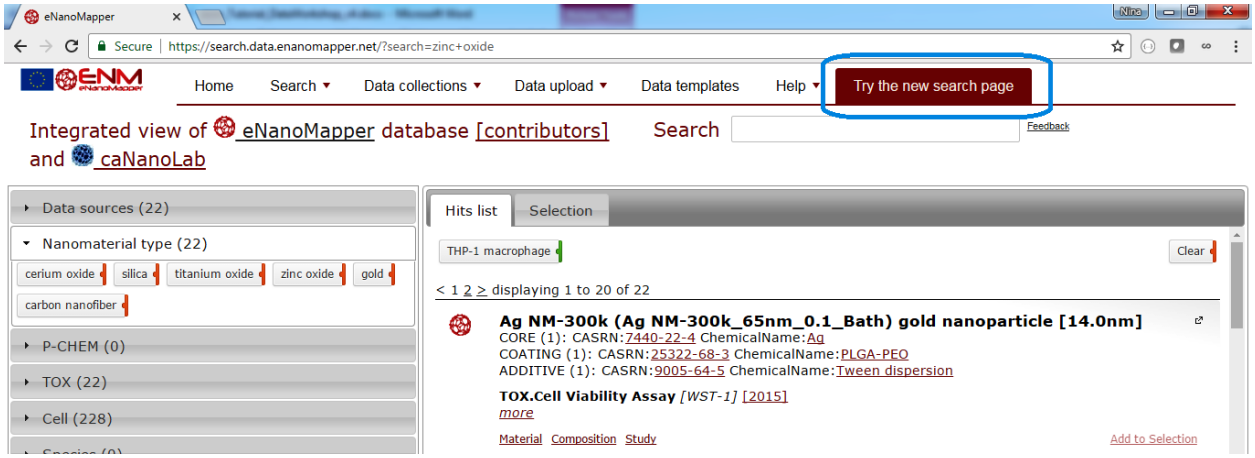


Figure 10 Use the menu at the top right to go to the new search page

Search for “carbon nanotube” (Fig 11). Clicking on each icon will lead to the study records in eNanoMapper or caNanoLab respectively. This is a demonstration of virtual search integration between the two databases, physically residing in servers in Europe and US, but with a common search interface.

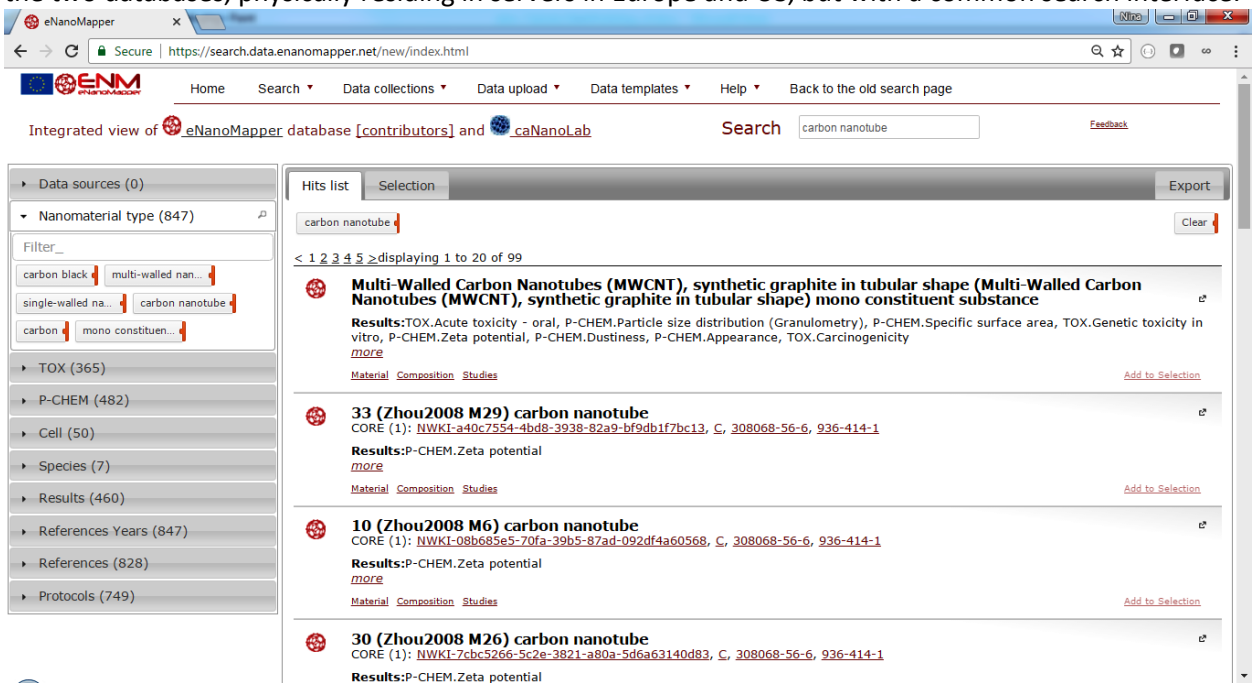


Figure 11 The screen shows results from both caNanoLab database (blue fullerene-like icon) and eNanoMapper database (red fullerene-like icon, starting from page 4). The “data sources” option at the left allow to further filter the results (e.g. NanoWiki is part of the eNanoMapper database).

The next exercise demonstrates how to filter the query to include results from single publication. Open the *References* widget at the left. (Figure 12)

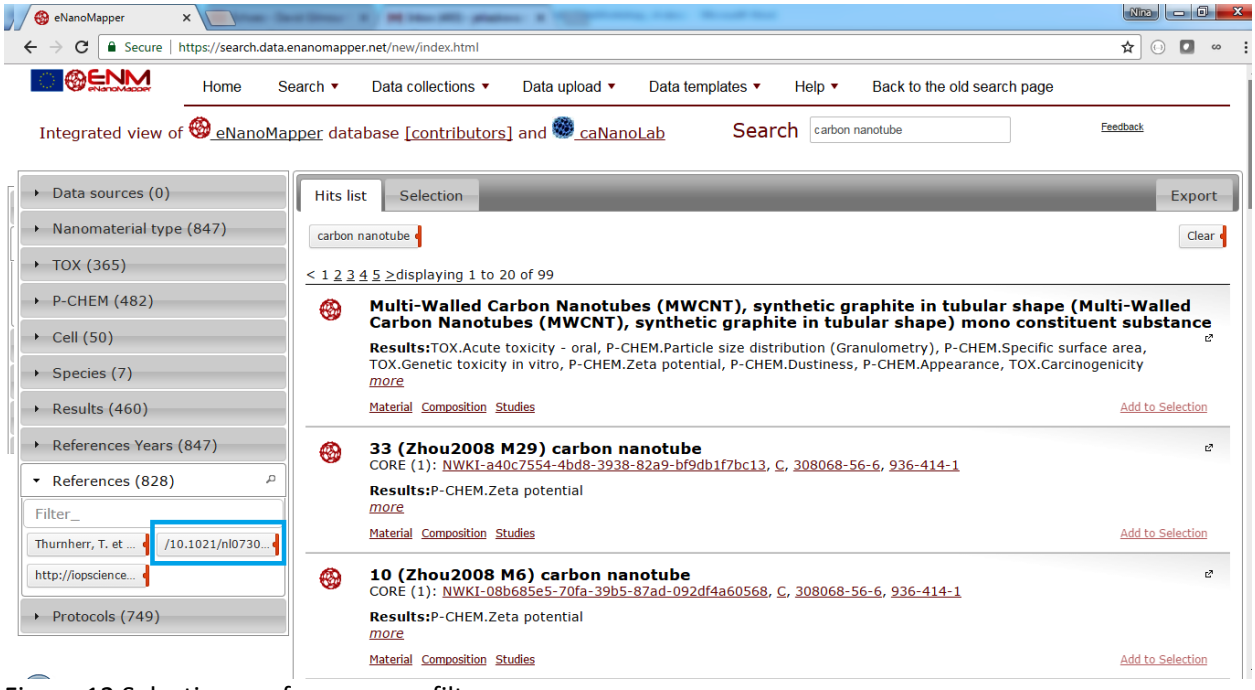
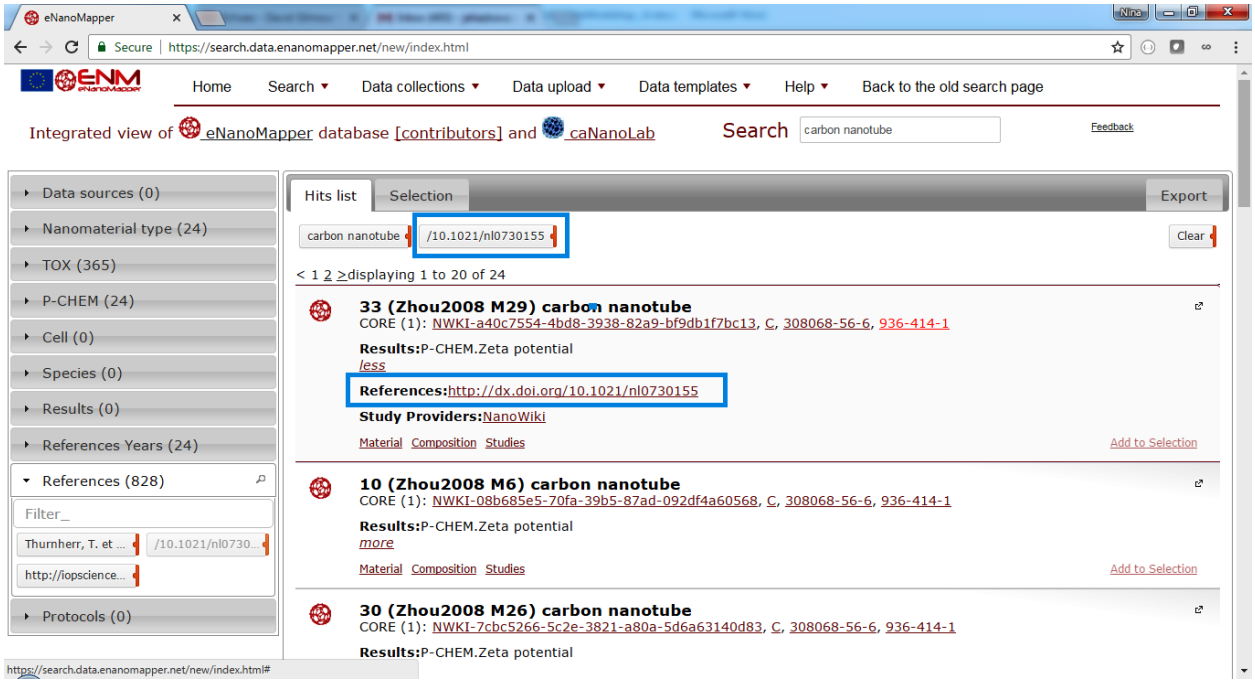
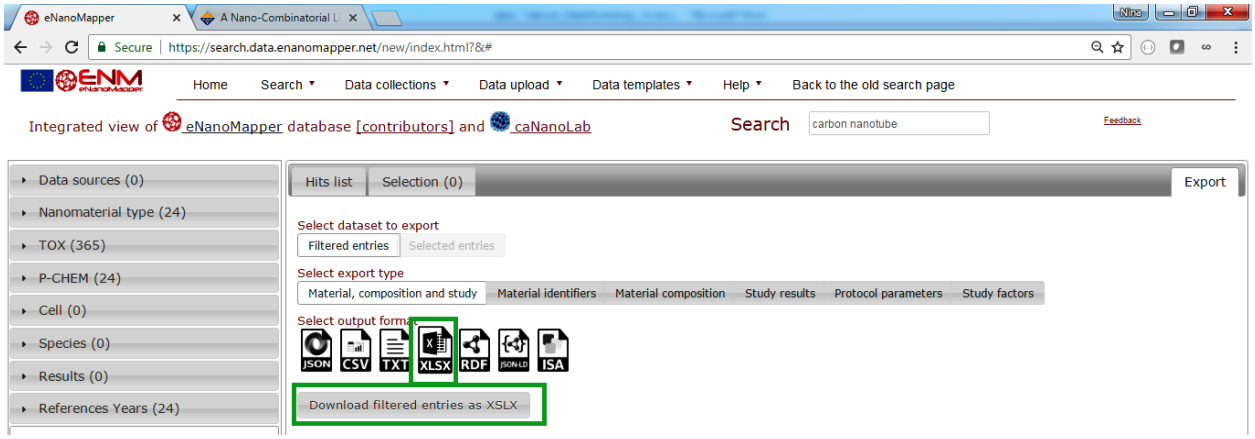


Figure 12 Selecting a reference as a filter

Click on the tag *10.1021/nl0730155* – it will be added to the current filter and the search hits restricted to materials from only this publication (Figure 13). Click on the link "more" to open more details about the material. A link to the publication page appears.



This version of the search application includes a new tab "Export" , allowing to download the query results in various formats. There are several options, defining what to download and the format. To sselect the Excel download, click on the XLSX icon and then click the "Download filtered entries as XLSX".



This completes the search and download exercise. You may explore other search and download options.

3. DOWNLOAD DATA FROM WITHIN R

This tasks introduces you the the R packages for accessing data in an eNanoMapper instance and will take about 15 mins. If you do not care about R, you can move on to the next chapter. Otherwise, follow the below steps.

3.1 INSTALL THE *RENM* PACKAGE

The *renm* package is an R package to download data from eNanoMapper data warehouse instances. It is available from: https://github.com/enanomapper/renm

From the R command line:

```
> install.packages(c("curl", "plyr", "jsonlite"))
> install.packages("devtools")
> library(devtools)
> install_github("enanomapper/renm")
```

The first line installs the dependencies of the *renm* package, while the second line installs a core R package that allows one to install packages directly from GitHub. The third line loads this *devtools* package, and installing of *renm* is finally done in the fourth line.

You can then load the renm library with:

```
> library(renm)
```

3.2 BROWSING THE DATA IN R

The eNanoMapper API (you can browse it interactively online) is wrapped by the *renm* package, hiding many technical aspects, attempting to adhere to R customs. For example, the following commands show you how to list all bundles or list all nanomaterials (modeled in the databases as substances):

```
> bundles = listBundles("http://data.enanomapper.net/")
> substances = listSubstances("http://data.enanomapper.net/")
```

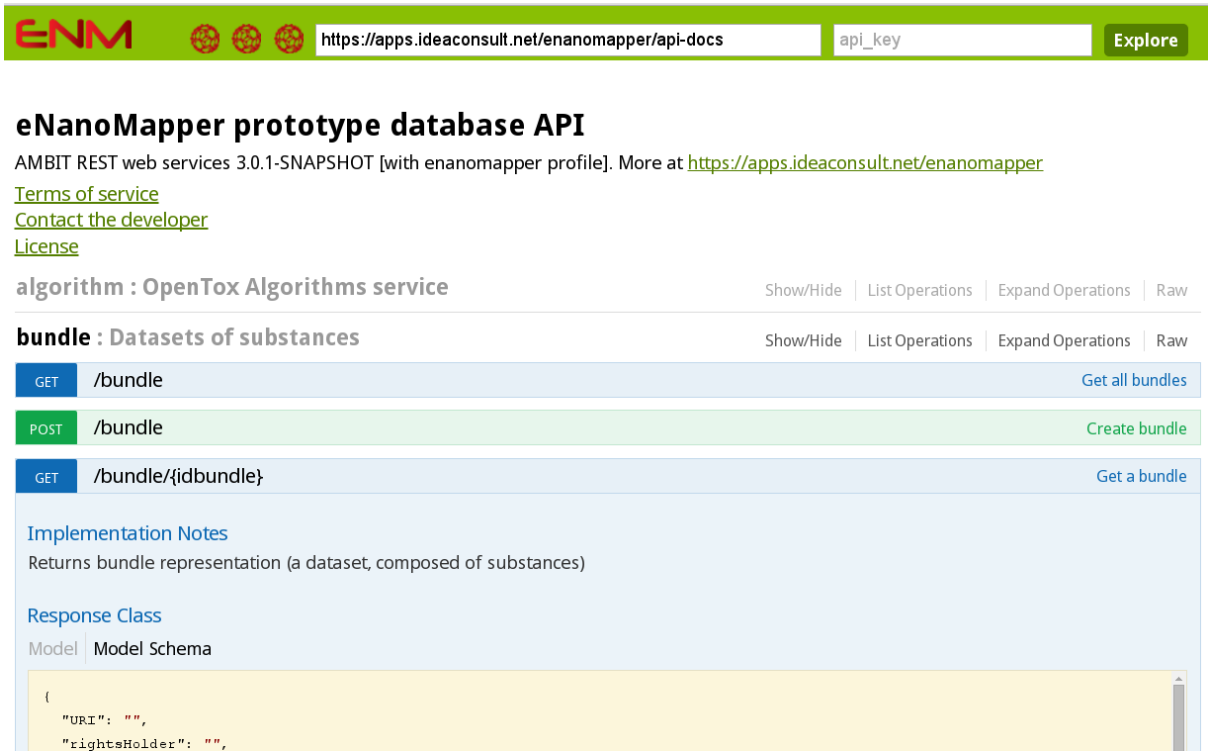


Figure 13 Screenshot of the webpage where you can explore and interactively try the application programming interface (API) use by the renm package for interaction with the data server.

2.3 DOWNLOAD THE DATA ON THE NANOMATERIALS

The following commands can be used to download information about nanomaterials.

```
> substances = listSubstances("http://data.enanomapper.net/")
> substanceFields = names(substances$substance)
> substanceLabels = substances$substance["name"]
```

The next step is to access the physchem and bioassay data associated with the nanomaterial:

```
> info =
substanceInfo("http://data.enanomapper.net/substance/NWKI-
71060af4-1613-35cf-95ee-2a039be0388a")
> experiments = info$protocol
```

And the get the first measurement:

```
> info$effects[[1]]
```

The structural information for the nanomaterials can be accessed in the following manner, but it's important to realize that here we see some design issues of the OpenTox API return, and the information we seek is expressed as 'features'. Furthermore, the content returned is automatically converted from the JSON returned by the service, and we need to extract too:

```
> structures = substanceStructures(
"http://data.enanomapper.net/substance/NWKI-71060af4-1613-35cf-95ee-
2a039be0388a"
)
```

```
> features = as.vector(unlist(
    lapply(structures$feature, function(aFeature) { aFeature$title })
))
```

The listSubstances() method can also be use to do searches. For example, to list all nanomaterials associated with a particular publication, you can use the Digital Object Identifier (DOI) of the paper:

```
> substances <- listSubstances(
    service="http://data.enanomapper.net/",
    search="10.1073/pnas.0802878105", type="citation"
)
```

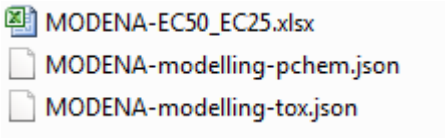
The list of types of searches currently supported include citation, type, and owner_name. The list of owners you can retrieve is in this way:

```
> owners = listSubstanceOwners("http://data.enanomapper.net/")
> codes = owners$value
```

4. DATA PREPARATION & UPLOAD

The next task will be about preparing data for upload into an eNanoMapper instance. You will look at a template with data, compiled by the MODENA COST Action project, explore the content, and upload this template into the test eNanoMapper data server https://apps.ideaconsult.net/enmtest .

The following three files are used in the exercise:



- *MODENA-EC50_EC25.xlsx* contains physicochemical and biological characterisation of two nanomaterials. The data was extracted from the literature, compiled by the MODENA project and made publicly available after the end of the MODENA project.
- *MODENA-modelling-pchem.json* is a configuration file, describing which parts of the Excel file *MODENA-EC50_EC25.xlsx* contain the physicochemical readouts and metadata of the experiment
- *MODENA-modelling-tox.json* is a configuration file, describing which parts of the Excel file *MODENA-EC50_EC25.xlsx* contain the biological readouts and metadata of the experiments
- The configuration files are written in JSON, a lightweight data interchange format.

4.1. DOWNLOAD THE FILES FROM THIS FOLDER

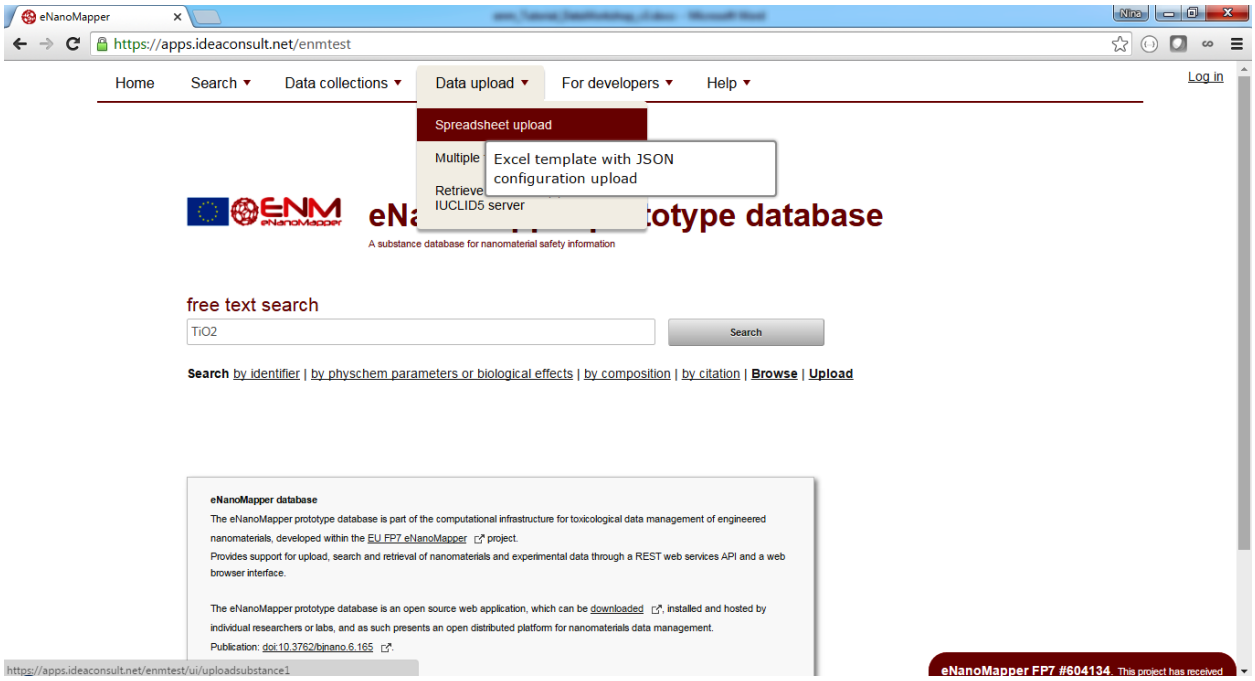
https://drive.google.com/drive/folders/0B55jmD17Vg55dVJqX3RXQ1JHMUk

4.2. GO TO HTTPS://APPS.IDEACONSULT.NET/ENMTEST

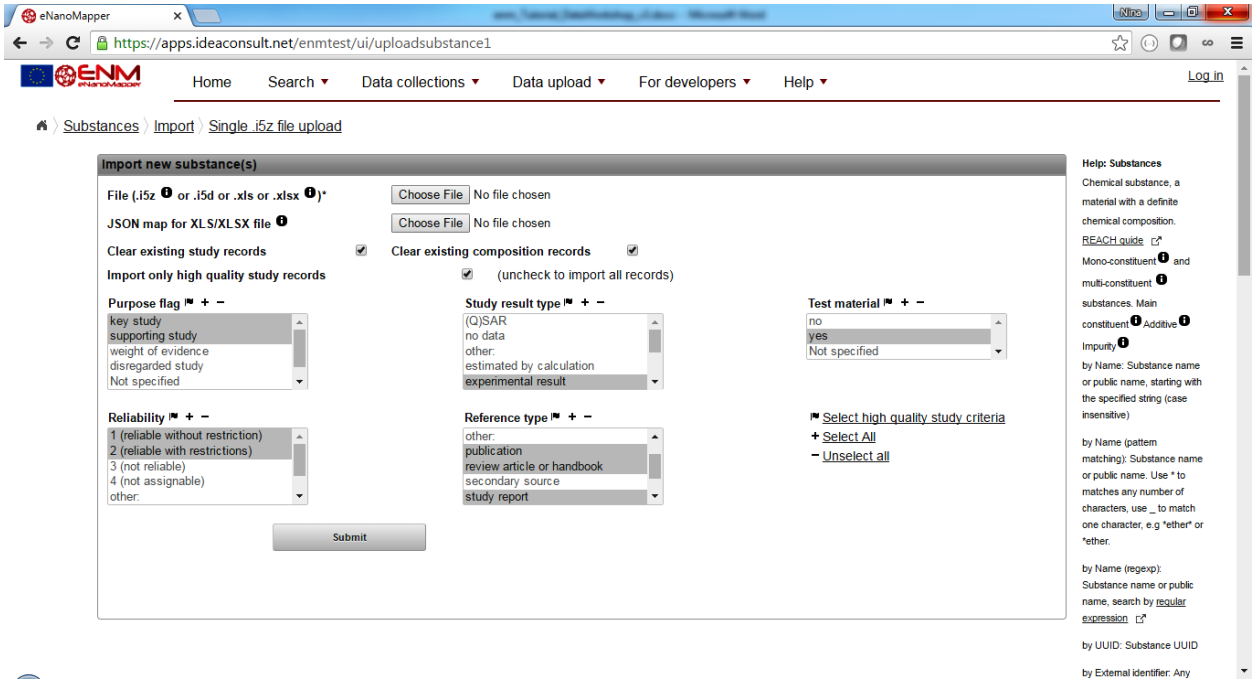
This is the test server for data upload.

4.3. USE THE MENU

DATA UPLOAD > SPREADSHEET UPLOAD

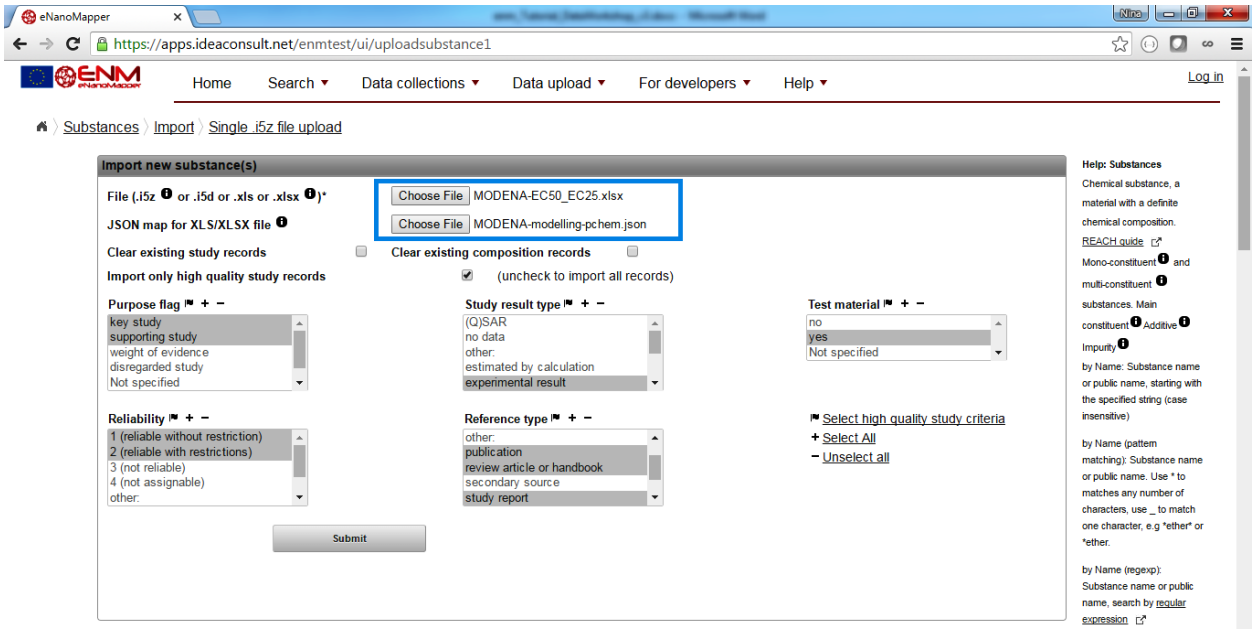


The following upload page will appear:

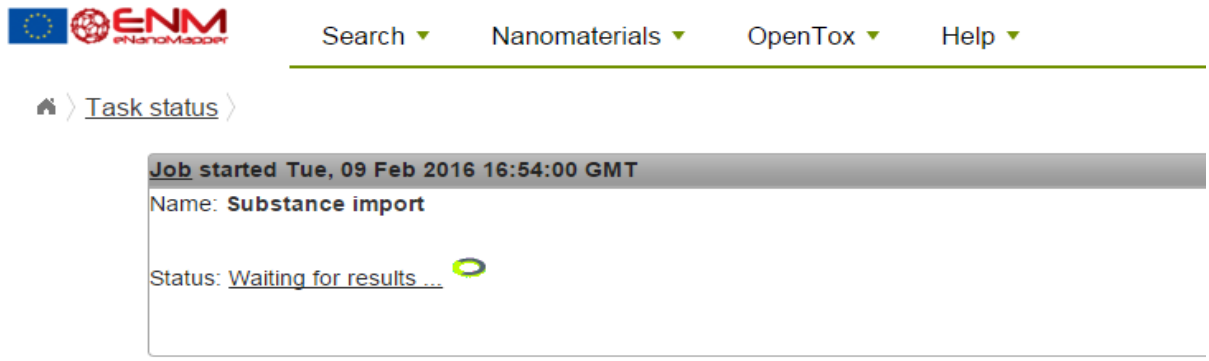


4.3. USING THE WEB FORM FOR UPLOAD

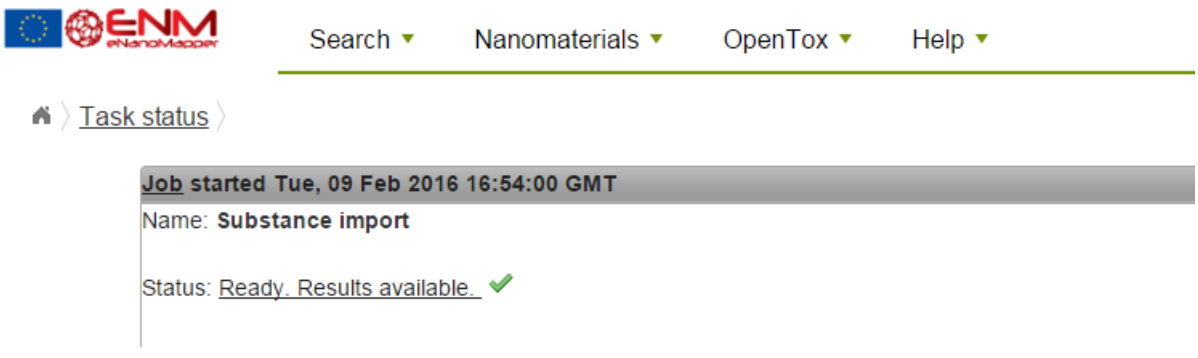
- Click on the top *Choose File* button and select the *MODENA-EC5_EC25.xlsx* file.
- Click on the bottom *Choose File* button and select the *MODENA-modelling-pchem.json*
- Uncheck the "*Clear existing study records*"
- Uncheck the *"Clear existing composition records"*
- Click *Submit*



Clicking the submission button starts the upload task, the following screen appears.

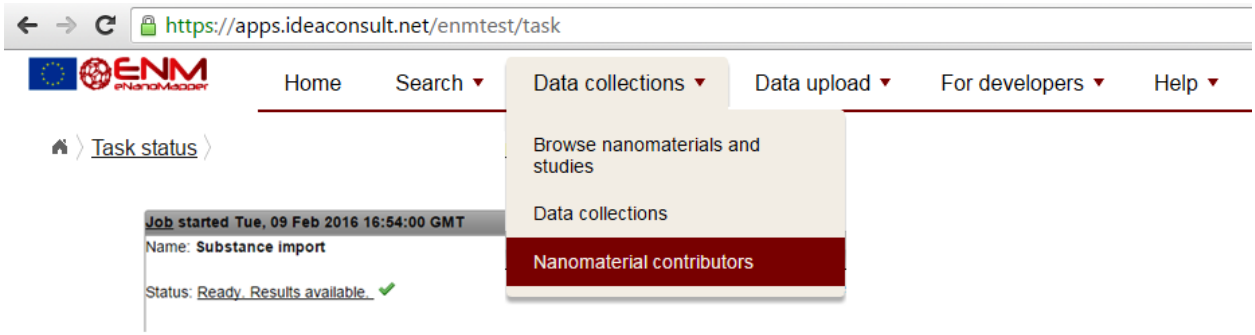


On completion, the status page changes to "*Ready. Results available*".

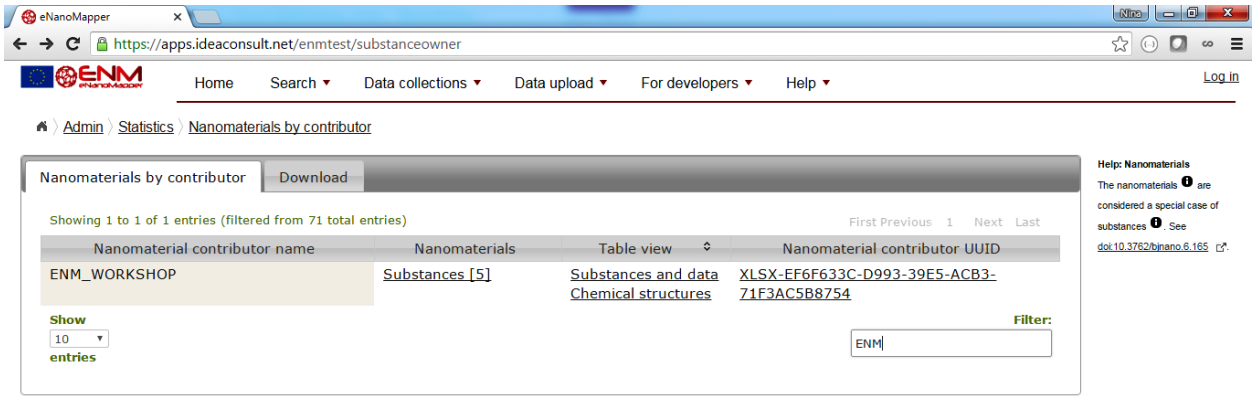


4.4. VIEW THE UPLOADED MATERIALS

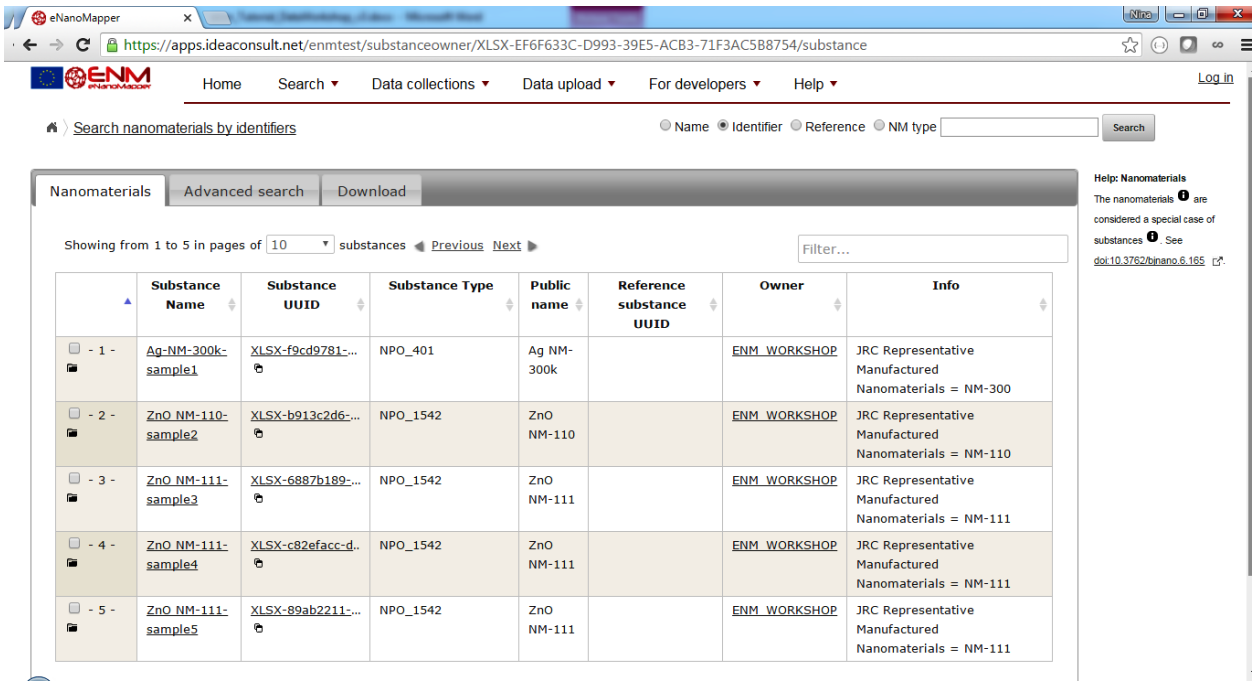
Clicking on the link will lead to the uploaded materials.
Alternatively, use the menu "*Data collections > Nanomaterial contributors*" to display the datasets.



- Look for the *Nanomaterial contributor name* ENM_WORKSHOP *(Hint: use the Filter function, marked in blue).*
- Click on the substances link to explore the uploaded content *(marked in red).*



The list of materials uploaded appears as shown in the next screenshot.



Each link in the columns *Substance name* leads to a detailed page with study details. Use the Expand all buttons to display and explore the results.

4.5. UPLOAD MODENA BIOLOGICAL CHARACTERISATION DATA

Repeat the upload with the step 4.3. and MODENA-modelling-tox.json in order to upload biological characterisation data.

Note that this exercise uses predefined files and configurations with the goal to get you familiar with the process of data upload. This exercise does not include data preparation task, such exercise will be subject of subsequent tutorials and / or bilateral interaction between eNanoMapper and NSC projects.

4.6. UPLOAD DOSE RESPONSE DATA

Download LDH assay data, generated by FP7 MARINA project (KI, [1]) from https://github.com/enanomapper/tutorials/tree/master/Hackathon_on_templates_for_data_collection/upload/blocks

There are two files, MARINA-invitro-WP09-P32-NM103TiO2-HMDM_Cytotoxity_LDH.xls and MARINA-invitro-config.json. The Excel file is a typical format used in many NanoSafetyCluster projects, and consists of several worksheets. Repeat the upload with the step 4.3. with these files.

[1] L. Farcal, F. Torres Andón, L. Di Cristo, B. M. Rotoli, O. Bussolati, E. Bergamaschi, A. Mech, N. B. Hartmann, K. Rasmussen, J. Riego-Sintes, J. Ponti, A. Kinsner-Ovaskainen, F. Rossi, A. Oomen, P. Bos, R. Chen, R. Bai, C. Chen, L. Rocks, N. Fulton, B. Ross, G. Hutchison, L. Tran, S. Mues, R. Ossig, J. Schnekenburger, L. Campagnolo, L. Vecchione, A. Pietroiusti, and B. Fadeel, "Comprehensive In Vitro Toxicity Testing of a Panel of Representative Oxide Nanomaterials: First Steps towards an Intelligent Testing Strategy," *PLoS One*, vol. 10, no. 5, p. e0127174, May 2015.

4.7. UPLOAD DATA IN JRC/NANOREG TEMPLATE (ADVANCED)

Download LDH template and the corresponding JSON configuration from https://github.com/enanomapper/tutorials/tree/master/Hackathon_on_templates_for_data_collection/upload/columns

There are two files, datatemplate_INVITRO_CYTOTOXICITY_LDH.xlsx and in_vitro_Cytotoxicity_LDH_sheet.json . The Excel file is part of the eNanoMapper release of data entry templates, based on JRC/NANoREG templates. More information at https://github.com/enanomapper/tutorials/tree/master/DataTemplates.

In this exercise the template is (almost) empty and before upload the data has to be entered. You may use the data from the file in 4.6 or enter your own data.

Repeat the upload with the step 4.3. with these files.

5. DATA VISUALIZATION

The last task of this tutorial shows you another advantage of Application Programming Interfaces: it is easy to reuse in different frameworks. Here, we show visualization of data with many HTML and JavaScript, using various libraries.

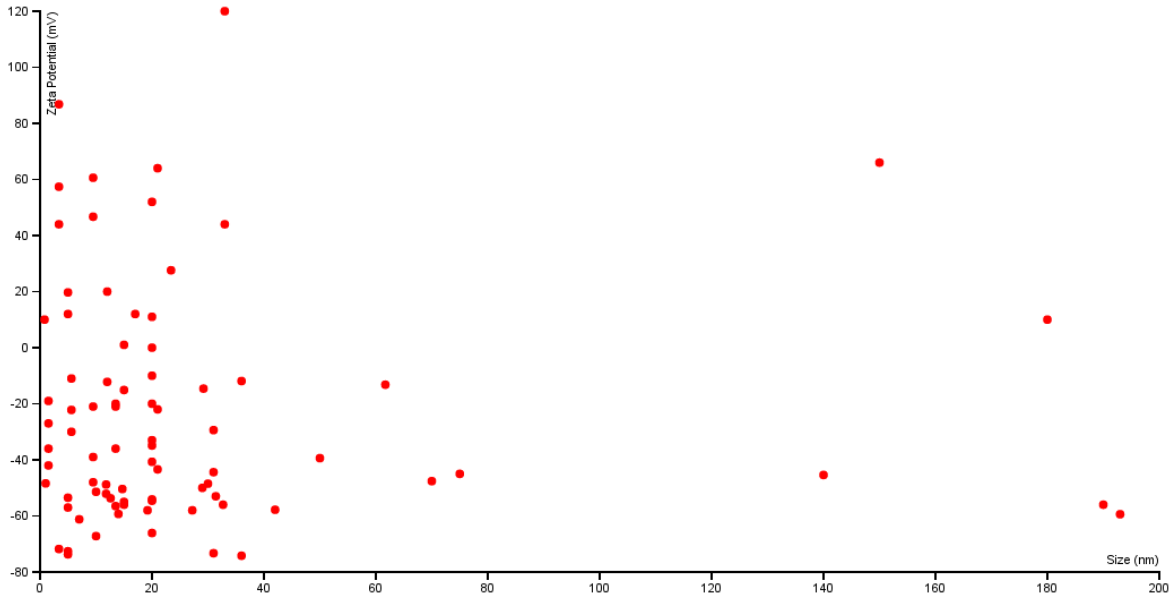
5.1 BROWSE THE EXAMPLES

Go to the following website and try the various examples. These JavaScript-enriched HTML pages use ambit.js and d3.js (mind you, these examples to not do justice to the features of d3.js!).

This page contains a few examples of HTML+JavaScript pages using the eNanoMapper API, using the ambit.js client (v0.0.2).

- example 1: show JSON response
- example 2: pie chart of substance by data sources
- example 3: pie chart of substance types (example 11: with labels)
- example 4: report the particle size of a specific substance
- example 5: histogram of material sizes of all substances
- example 6: histogram of zeta potentials of all substances (at any pH)
- example 7: list all substances in the database from the Beilstein Journal of Nanotechnology
- example 8: list all substances in the database from the ACS journals
- example 9: list data bundles
- example 10: scatter plot showing zeta potential versus material size (no correlation)

Particularly, check example 5, example 9, example 10, and example 11.



5.2 TITANIUM DIOXIDE

Open this page to view a summary of titanium dioxides in data.enanomapper.net:
http://enanomapper.github.io/ambit.js/titaniumoxide.html