DELIVERABLE REPORT D5.7

D5.7 Final report on User Applications

GRANT AGREEMENT:	604134
ACRONYM:	eNanoMapper
NAME:	eNanoMapper - A Database and Ontology Framework for Nanomaterials Design and Safety Assessment
PROJECT COORDINATOR:	Douglas Connect GmbH
START DATE OF PROJECT; DURATION:	1 February 2014; 36 months
PARTNER(s) RESPONSIBLE FOR THIS DELIVERABLE:	DC
DATE:	31.10.2016
VERSION:	V.1.1

Call identifier	FP7-NMP-2013-SMALL-7
Document Type	Deliverable Report
WP/Task	WP5 – Tasks T5.3 and T5.9
Document ID	eNanoMapper D5.7 Final report on User Applications
Status	Draft

Partner Organisations	• Douglas Connect, GmbH (DC) • National Technical University of Athens (NTUA) • In Silico Toxicology (IST) • Ideaconsult (IDEA) • Karolinska Institutet (KI) • European Bioinformatics Institute (EMBL-EBI) • Maastricht University (UM) • Misvik Biology (MB)
Authors	Lucian Farcal, Philip Doganis, Georgia Tsiliki, Haralambos Sarimveis, Nina Jeliazkova, Egon Willighagen, Linda Rieswijk, Penny Nymark, Micha Rautenberg, Christoph Helma, Maja Brajnik, Barry Hardy
Purpose of the Document	To report on the implemented user applications and Task 5.3 and 5.9 related activities
Document History	1. First draft, 26/09/2016 2. Second draft 13/10/2016 3. Final version, 04/11/2016

TABLE OF CONTENTS

TABLE OF FIGURES

1. EXECUTIVE SUMMARY

Deliverable 5.7 reports on the Tasks 5.3 and 5.9 on the user applications achievements. The application infrastructure developed within eNanoMapper project aims to support the data management in the area of nanosafety research and to enable an integrated approach for the risk assessment of nanomaterials. To achieve these, eNanoMapper developed an ontology, a data infrastructure and modelling tools with applicability in risk assessment of nanomaterials. eNanoMapper developed resources and tools for predicting toxicity of nanomaterials and worked towards improving the standards in risk assessment of nanomaterials.

The testing phase aimed to provide the basis for test-driven development, ensure the interoperability of services and the developing ontology and alert developers about bugs and interoperability problems. Moreover, the tools and functionalities development followed as much as possible the prioritized use cases identified within WP1.

- The ontology includes common vocabulary terms used in nanosafety research and aims to provide a clear explanation of nanostructures based on information relating to their characterization, relevant experimental paradigms, biological interactions, safety indications and the integration of data from existing nanotoxicology sources.
- To support a collaborative safety assessment approach, an infrastructure for data management was developed, with a database which includes functionalities for data protection, data sharing, data quality assurance, search and interfaces for different needs and usages, comparability and cross-talk with other databases.
- In addition, a collection of descriptors, computational toxicology models and modelling tools were developed, to enable the use and integration of nanosafety data from various sources, via nanoQSAR modelling, read across prediction modelling, experimental design, and interlaboratory testing.

The project provides also a rich source of information and documentation (e.g. tutorials, webinars, publications) to support and guide the users. eNanoMapper dictionary of terms is also available at: http://www.enanomapper.net/library/enm-dictionary

Figure 1. eNanoMapper portal for user applications (http://www.enanomapper.net/applications)

2. INTRODUCTION

Deliverable 5.7 aims to report on the Tasks 5.3. and Task 5.9 on the user applications achievements (technical options, user interfaces) and link with the prioritized use cases previously identified and described (*see Deliverable 1.2*). The applications followed different steps from the development and optimization phase to testing, before being released for user testing and evaluation. A detailed documentation was recorded for each application.

Within Task 5.3 Integrated Testing, the testing processes were defined and executed, on the prototype or released versions of the applications. The testing aimed to evaluate all applications and it was performed individually or jointly (e.g. within hackathon sessions dedicated to specific applications). Guidance materials were prepared and feedback sections were opened for capturing eventual issues from the users.

The testing phase aimed to provide the basis for test-driven development, ensure the interoperability of services and the developing ontology and alert developers about bugs and interoperability problems.

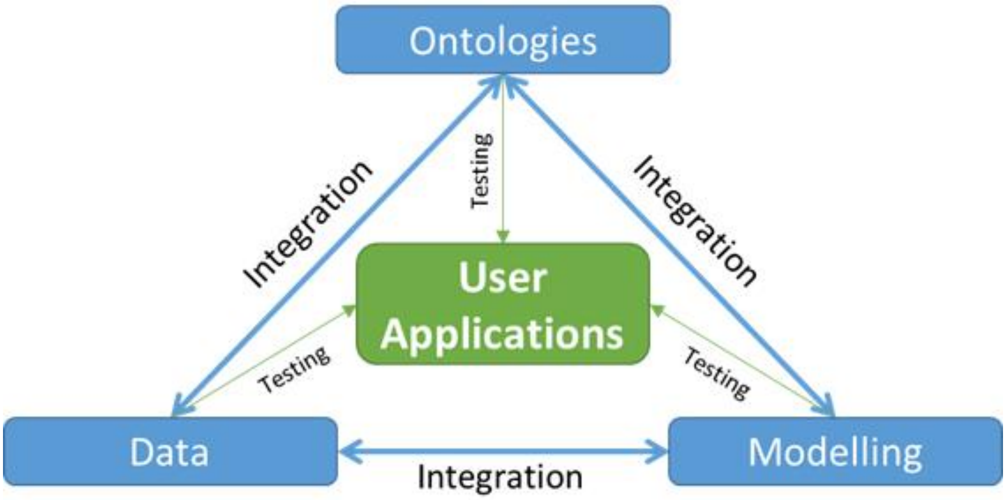


Figure 2. Integration of ontologies, data and modelling tools towards developing user applications

The aim of Task 5.9 User Application for Prioritised Use Cases, was to guide the development of the user applications based on the case studies identified at the beginning of the project within WP1. As shown in Chapter 4 of this report, most of the prioritised use cases are linked to the user applications, which cover one or more of these cases:

- A comprehensive **ontology** was developed for the nanosafety domain, encompassing nanomaterials and all information relating to their characterization, as well as information describing relevant experimental paradigms, biological interactions, and safety indications. The eNanoMapper ontology includes common vocabulary terms in use in nanosafety research with a classification hierarchy and other relationships. The ontology includes a refinement of the NanoParticle Ontology and incorporates terms from other high quality ontologies (e.g. OBI, ENVO etc.).

- A **data infrastructure and search capabilities** was developed, supporting all aspects of nanomaterials characterization, environmental and health hazards assessment, and high throughput and high content datasets. The eNanoMapper database is a public database hosting nanomaterials composition and characterization data and biological and toxicological information. The database software is open source, and download / installation instructions are available. The database provides various possibilities to search and explore information, and to download data in different formats as a web application or service.

  The import and upload functionality currently supports a variety of NSC templates and OECD Harmonized Templates. The importer is easily extensible to support more spreadsheet formats through configuration.

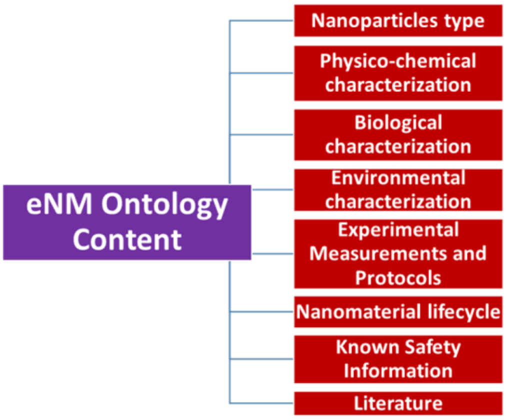
- **Modelling infrastructure and services** were developed to analyse nanomaterials-related descriptors and experimental data. The online platform Jaqpot Quattro uses data available from data.enanomapper.net and allows building predictive nanoQSAR models enriched by multiple validation schemes, as well as predictive read across models. Additional functionalities include experimental design and interlaboratory testing techniques. Datasets can be preprocessed or undergo PMML transformations before modelling. Users can upload their own models or algorithms to the system, or seamlessly integrate models and algorithms hosted elsewhere. A user-friendly interface is now available that enables access to eNanoMapper Jaqpot Quattro tools and applications, but also to a repository of validated models. Other developed applications are also available, i.e. the Image descriptor Calculation Tool, a systematic framework for the automated analysis of microscopy images of nanomaterials and the calculation of morphology and population-related nanoparticle descriptors, the RRegrs R package, a collection of R regression tools aiming to find optimal and well validated nanoQSAR models with applications to chemoinformatics and nanotoxicology data, and the Gene Ontology descriptors tool, a framework to group omics NP data to descriptors that would easily summarize the biological data.

- The modelling service links the eNM database with the modelling platform, e.g.:

- It transforms experimental data into a modelling-friendly format, and produces standardized datasets.
- Raw data of nanomaterials such as images, crystallographic data and proteomics data, are processed by descriptor calculation services to produce compact and informative numerical descriptors.
- Data modelling and analysis services implement computational experimental design methods, tools for interlaboratory proficiency testing and non-testing approaches for predicting toxicity and for filling data gaps (NanoQSAR and read-across methods), offering the option of evaluating a model’s accuracy and robustness.
- A number of mechanistic modelling tools combine experimental data with information from public biological databases and tools to perform pathway and network analysis.

3. USER APPLICATIONS

3.1. ONTOLOGY



An ontology is a controlled vocabulary built on the hierarchical relationships between terms, used, for example, to harmonize data from several sources to one database or analyze the relationships between terms and categories of terms.
The eNanoMapper ontology is developed by the eNanoMapper team. The ontology includes and defines common vocabulary terms in use in nanosafety research with a classification hierarchy and other relationships. In parts it is an extension and refinement of the NanoParticle Ontology.

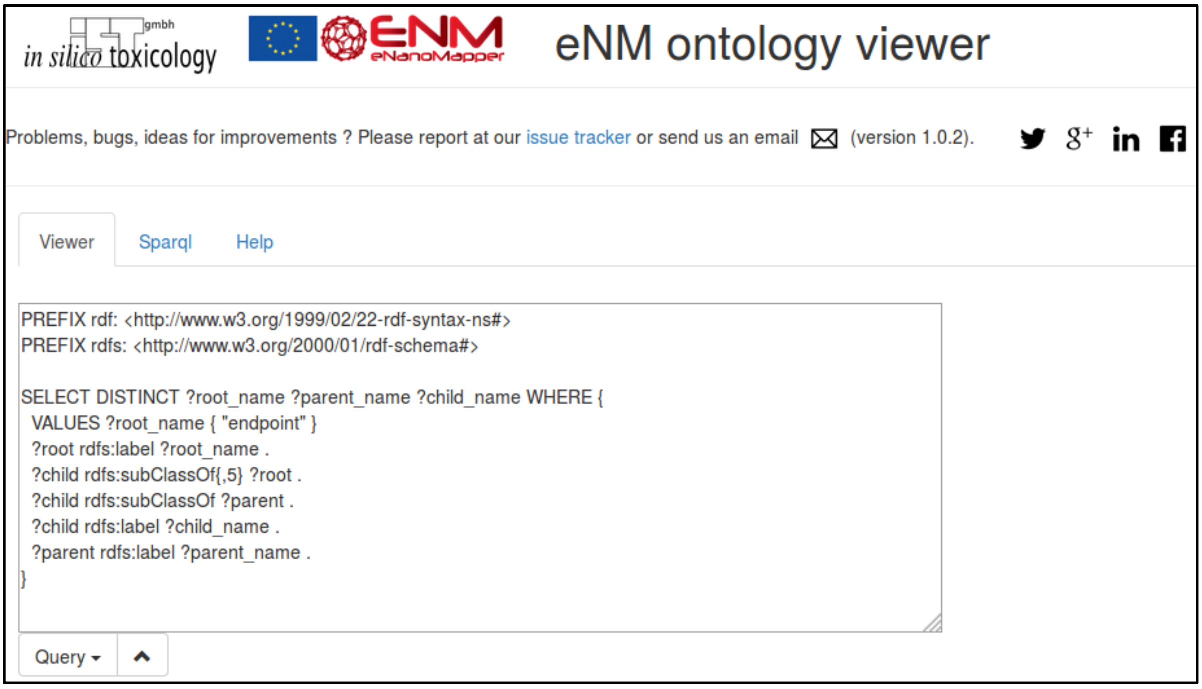
- Download the ontology in .owl from: http://purl.enanomapper.net/onto/enanomapper.owl
- Download and explore on Bioportal: http://bioportal.bioontology.org/ontologies/ENM
- Download and explore on Aber-OWL: http://aber-owl.net/ontology/ENM
- Download and explore on Ontology Lookup Service: http://www.ebi.ac.uk/ols/beta/ontologies/enm
- Feedback: https://github.com/enanomapper/ontologies/issues

3.1.1. ONTOLOGY VIEWER APPLICATIONS

3.1.1.1 ENANOMAPPER ONTOLOGY VIEWER

Web application and graphical interface for the eNanoMapper triple store:

- Viewer: https://purl.enanomapper.net/enm-viewer
- Development: https://github.com/enanomapper/enm-ontoviewer
- Feedback: https://github.com/enanomapper/enm-ontoviewer/issues
- Technology: Web application, Library / Ruby, JavaScript
- License / Waiver: GPL3



Problems, bugs, ideas for improvements ? Please report at our issue tracker or send us an email (version 1.0.2).

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

SELECT DISTINCT ?root_name ?parent_name ?child_name WHERE {
 VALUES ?root_name { "endpoint" }
 ?root rdfs:label ?root_name .
 ?child rdfs:subClassOf{,5} ?root .
 ?child rdfs:subClassOf ?parent .
 ?child rdfs:label ?child_name .
 ?parent rdfs:label ?parent_name .
}
```

Figure 3. eNM ontology viewer application (search interface)

The eNanoMapper ontology viewer application (Figure 3) visualises SPARQL query results from eNM ontologies and stored in a triple-store database[1]. Data in the triplestore is mirrored in RDF format from the eNanoMapper ontology and database (https://data.enanomapper.net/). The eNanoMapper ontology viewer application is publicly available at http://purl.enanomapper.org/enm-viewer. Details about graphics and file output options with screenshots can be found in Annex 6.2.

3.1.1.2. ONTOLOGY VISUALISATION

Access: https://search.data.enanomapper.net/treedemo/

Study summary statistics:
- data.enanomapper.net
- ambitlri.ideaconsult.net

Ontology visualisation:
- Chemical substance (CHEBI59999)
- eNanoMapper ontology
- eNanoMapper ontology (subclasses of BFO_0000040)
- BioAssay ontology (BAO)
- Gene ontology (GO)

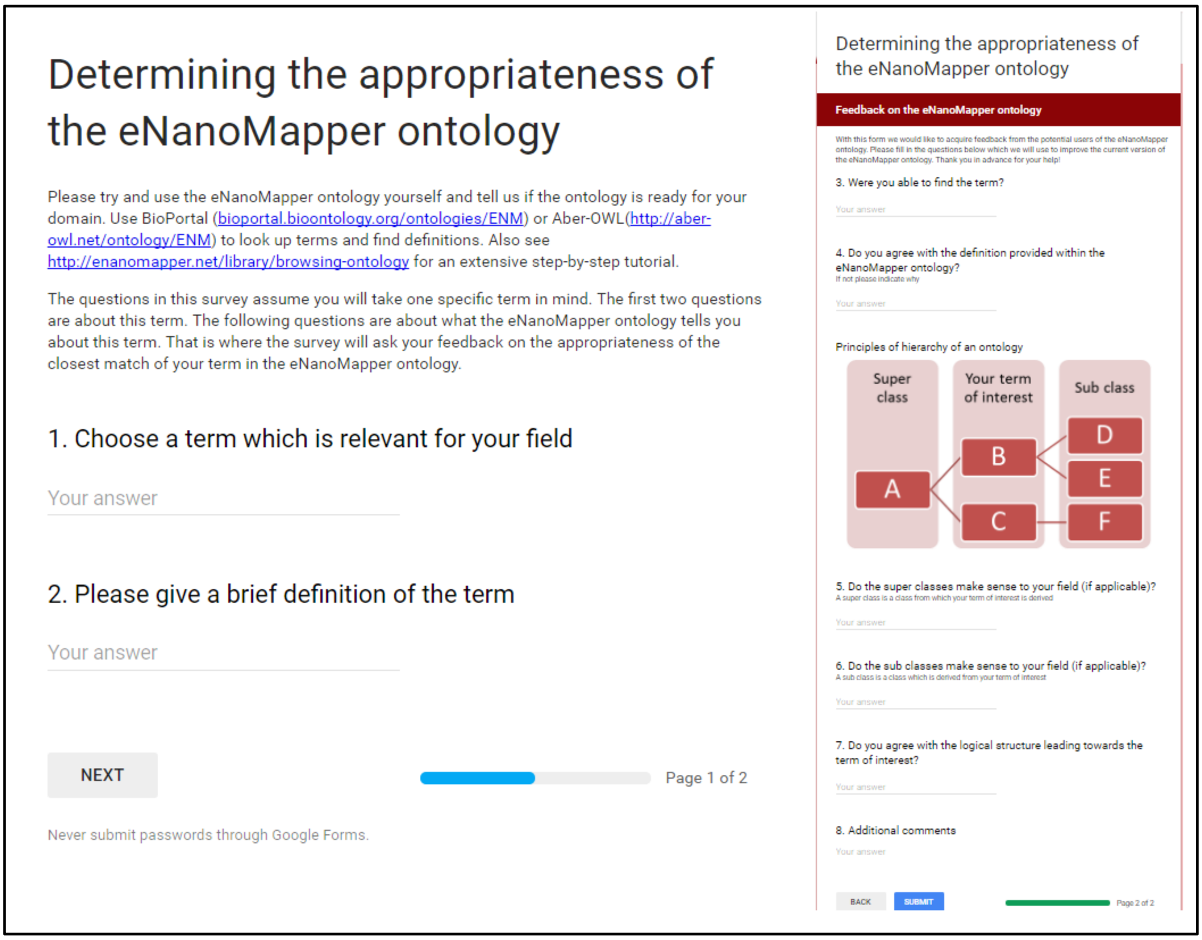
Feedback: https://github.com/enanomapper/data.enanomapper.net/issues/59

---

[1] The SPARQL endpoint triplestore is described in Deliverable D3.3

3.1.2. USER EVALUATION OF ONTOLOGY

In order to determine the appropriateness of the ENM ontology, a survey was created using Google Forms. The survey focuses on evaluating the ontology for users within the nanosafety domain. The eight questions within the survey asks users to access the ENM ontology online either via BioPortal or using Aber-OWL and explore its use with keeping example terms in mind. It also refers to the tutorial which has been created to browse the eNM ontology. When the user has been looking for the example terms in the eNM ontology they are asked about the appropriateness of the closest match of the example term in the eNM ontology. The outcomes of the survey will be used for the external evaluation of the ENM ontology. Screenshots of the survey are listed in Figure 4.



Determining the appropriateness of the eNanoMapper ontology

Please try and use the eNanoMapper ontology yourself and tell us if the ontology is ready for your domain. Use BioPortal (bioportal.bioontology.org/ontologies/ENM) or Aber-OWL(http://aber-owl.net/ontology/ENM) to look up terms and find definitions. Also see http://enanomapper.net/library/browsing-ontology for an extensive step-by-step tutorial.

The questions in this survey assume you will take one specific term in mind. The first two questions are about this term. The following questions are about what the eNanoMapper ontology tells you about this term. That is where the survey will ask your feedback on the appropriateness of the closest match of your term in the eNanoMapper ontology.

1. Choose a term which is relevant for your field

Your answer

2. Please give a brief definition of the term

Your answer

Page 1 of 2

Never submit passwords through Google Forms.

Determining the appropriateness of the eNanoMapper ontology

Feedback on the eNanoMapper ontology

With this form we would like to acquire feedback from the potential users of the eNanoMapper ontology. Please fill in the questions below which we will use to improve the current version of the eNanoMapper ontology. Thank you in advance for your help!

3. Were you able to find the term?

Your answer

4. Do you agree with the definition provided within the eNanoMapper ontology?
If not please indicate why

Your answer

Principles of hierarchy of an ontology

5. Do the super classes make sense to your field (if applicable)?
A super class is a class from which your term of interest is derived

Your answer

6. Do the sub classes make sense to your field (if applicable)?
A sub class is a class which is derived from your term of interest

Your answer

7. Do you agree with the logical structure leading towards the term of interest?

Your answer

8. Additional comments

Your answer

Page 2 of 2

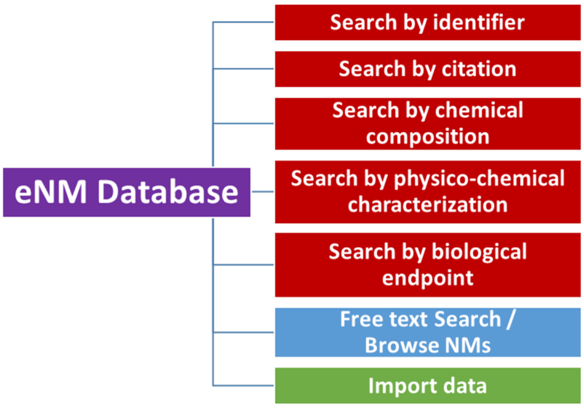
Figure 4. The survey to evaluate the appropriateness of the eNM ontology

3.1.3. DOCUMENTATION ON ONTOLOGY

~ Ontology

- Development
    - ~ NanoParticle Ontology
    - ~ Ontologies
        - ~ Feedback
        - ~ Licences
        - ~ Basic Formal Ontology
    - ~ Slimmer
    - ~ Test suite
- Applications
    - ~ Download .owl
    - ~ Bioportal
    - ~ Aber-OWL
    - ~ Ontology Lookup Service
    - ~ Ontology Viewer
    - ~ Ontology Visualisation
- User guidance
    - Tutorials
        - ~ Browsing the eNM ontology with bioportal, aberowl and protege
    - Webinars
        - ~ Using the eNanoMapper ontology
    - Publications
        - ~ Hastings et al., 2015, Journal of Biomedical Semantics
        - ~ Hastings et al., 2014, JBMS supplement: Bio-Ontologies meeting
- ~ DoW
    - ~ Issue tracker
    - Reporting
        - ~ D2.1 Framework and Infrastructure for Ontology development, versioning and dissemination
        - ~ D2.2 Ontology Content Types and Existing Community efforts
        - ~ D2.3 Ontology initial release
        - D2.4 Ontology final release

3.2. DATA MANAGEMENT APPLICATIONS



3.2.1. ENANOMAPPER DATABASE DEMO

data.eNanoMapper.net is a public database demo hosting nanomaterials characterization data and biological and toxicological information. The database provides various possibilities to search and explore information, and to download data in various standard formats. The database supports data upload in various formats, including configurable Excel templates. The server contains data from various sources.

- Database technologies available: SQL databases; supported formats: IUCLID5 , ISA-TAB, CSV, Excel spreadsheet, RDF, JSON
- Types of NMs involved: various
- Link: data.enanomapper.net
- Source and download: http://ambit.sourceforge.net/enanomapper.html
- License / Waiver: various Open Source licenses

3.2.2. SEARCH APPLICATION

The search engine integrates data.enanomapper.net and caNanoLab.

- Access: https://search.data.enanomapper.net/
- Search application dedicated to NANoREG data (*under development*)
  https://sandbox.ideaconsult.net/search/nanoreg1

3.2.3 PROTOCOL SERVICE

The protocol service is available at https://apps.ideaconsult.net/enmprotocol
Source code: https://github.com/enanomapper/toxbank-api-server

3.2.4. DATA MANAGEMENT DEVELOPMENT TOOLS

3.2.4.1. ENANOMAPPER TRIPLESTORE AND SPARQL ENDPOINT

Aggregation of eNanoMapper data and ontologies in RDF for SPARQL queries. A virtuoso RDF triplestore wrapped in a ruby webservice with graphical interface and machine accessible SPARQL endpoint.

- Access: https://purl.enanomapper.net/enm-viewer
- Endpoint: https://purl.enanomapper.net/sparql
- Technology: Virtuoso / Ruby

3.2.4.2. AMBIT.JS

JavaScript client library for the eNanoMapper / AMBIT API.

- Status: beta
- Technology: Library / JavaScript
- Access: https://github.com/enanomapper/ambit.js/
- Demo: http://enanomapper.github.io/ambit.js/
- License: MIT License

3.2.4.3. NM DATAPARSER

A command line, which can convert different formats

3.2.4.4. NANO-LAZAR ENM PARSER

lazar (https://github.com/opentox/lazar) comes with command line tools and scripts to mirror, parse and import eNanoMapper data, based on the eNM REST API.

3.2.4.5. R ENANOMAPPER CLIENT PACKAGE

R client library for the eNanoMapper / AMBIT API.

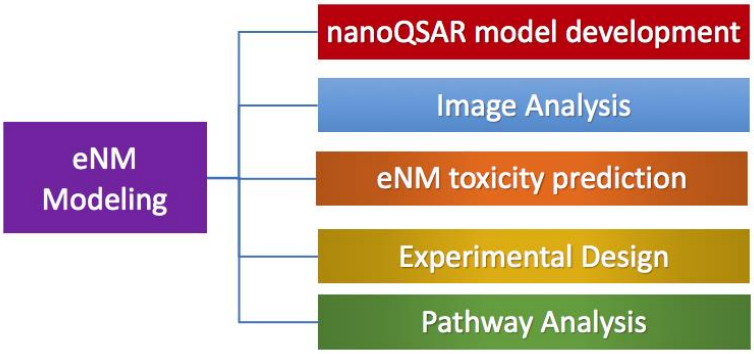
- Status: beta
- Technology: Library / R
- Access: https://github.com/enanomapper/renm/
- Demo: http://enanomapper.github.io/ambit.js/
- License: MIT License

3.2.5. DOCUMENTATION ON DATA MANAGEMENT APPLICATIONS

~ Data

- Development
    - ~ API
    - ~ AMBIT web services
    - ~ Triplestore and SPARQL endpoint
    - ~ AMBIT.JS
    - NM Dataparser
    - ~ Nano-lazar parser
    - ~ R eNanoMapper client package
- Applications
    - Database
        - ~ Data access
        - ~ Search
        - ~ Data sources
        - ~ Protocol Service
- User guidance
    - Tutorials
        - ~ Entering and analysing nanosafety data
        - ~ AMBIT user guide
    - Webinars
        - ~ A substance database to support a safe innovation approach
        - ~ Data visualization with AMBIT.JS and D3JS
    - Publications
        - ~ Jeliazkova et al., 2015, Beilstein Journal of Nanotechnology
        - ~ Jeliazkova et al., 2014, Bioinformatics and Biomedicine (BIBM)
- ~ DoW
    - ~ Issue tracker
    - Reporting
        - ~ D3.1 Technical Specification and initial implementation of the protocol and data management web services
        - ~ D3.2 Data Management System with extended search capabilities
        - D3.3 Modules and services for linking and integration with third party databases
        - D3.4 ISA-Tab templates for common bioselected set of assays

3.3. MODELLING APPLICATIONS AND DESCRIPTORS



3.3.1. JAQPOT QUATTRO

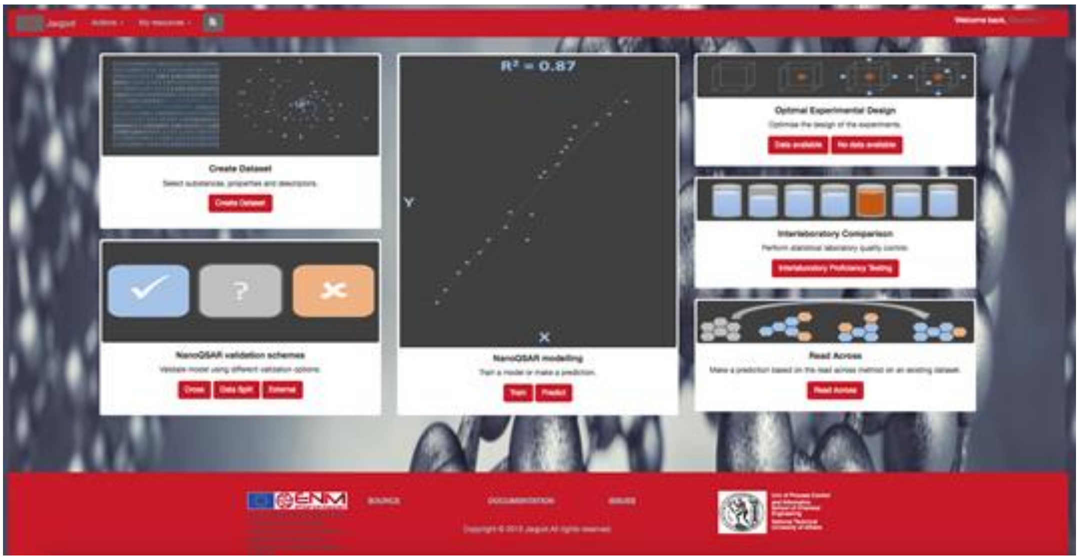


Figure 5. JaqPot Quattro user interface

JaqPot Quattro can be considered as an extendable suite of computational tools under a common framework (see also Deliverable 4.3, Deliverable 4.5, Deliverable 4.6). Current functionalities include:

- Data preprocessing algorithms
- Computing descriptors from images and crystallographic and omics data
- Integrating diverse data and creation of datasets suitable for modelling
- Training and validation of nanoQSAR models using a rich selection of machine learning and data mining algorithms (which will be further extended)
- Predictive modelling using the read across method
- Repository of predictive models as ready to use web applications
- Integration with R, Python, Weka and access to JaqPot services through Knime
- Factorial and Optimal experimental design
- Interlaboratory testing implementing ISO 13528
- Reporting results in automatically created and downloadable pdf files (including QPRF OECD reports)
- Saving models in PMML format for easy transfer to other platforms

- Allowing users to create and store in the cloud their own resources (datasets, models, reports) which they can access and manage at any time and place through a web browser
- Authorisation and authentication services
  - Architecture of tool / Technology: REpresentational State Transfer (REST) / Java, Python, R
  - API: http://app.jaqpot.org:8080/jaqpot/swagger/
  - Source code: https://github.com/KinkyDesign/JaqpotQuattro
  - Access / download: http://www.jaqpot.org
  - License/waiver: GPL3

3.3.2. IMAGE DESCRIPTORS CALCULATION TOOL

Tool for automated calculation of descriptors from images of nanomaterials. The tool provides the user with a systematic framework for the automated analysis of microscopy images of nanomaterials and the calculation of nanoparticle descriptors (see also Deliverable 4.2).

- Architecture of tool / Technology: Web application / Java
- Source code: https://github.com/enanomapper/imageAnalysis
- Access: http://app.jaqpot.org:8880/imageAnalysis/
- License/waiver: GPL3

3.3.3. RREGRS

RRegrs is an R package for computer-aided model selection for multiple regression methodologies by means of an extensive cross-validation scheme. It is used to find the best regression models for any numerical dataset. The main use of the script is aimed at finding optimal and well validated QSAR models for chemoinformatics and nanotoxicology.

- Architecture of tool / Technology: R
- Source code, access, download: https://github.com/enanomapper/RRegrs
- License/waiver: MIT / BSD

3.3.4. TOXFLOW

toxFlow is a shiny web application for enrichment analysis and predictive read across modelling. If physicochemical and omics NP data are available both techniques can be applied in a sequential order. The tool has been developed by Dimitra D. Varsou during her undergraduate degree project.

- Architecture of tool / Technology: R
- Source code, access, download: http://147.102.86.129:3838/
- License/waiver: GPL3

3.3.5. CHIPSTER FOR NANOMATERIAL-BASED OMICS DATA ANALYSIS

Chipster is an open source user-friendly analysis software for high-throughput data analysis. It offers over 350 bioinformatics tools and it is constantly updated according to the latest state-of-the-art tools and scripts. Users can analyse and visualize data interactively, and share complete analysis sessions and automatic workflows with colleagues. Chipster supports the analysis of over 120 different microarray platforms, including the most common types by Affymetrix, Agilent and Illumina. In addition, Chipster has extensive functions and tools for the analysis of NGS data.

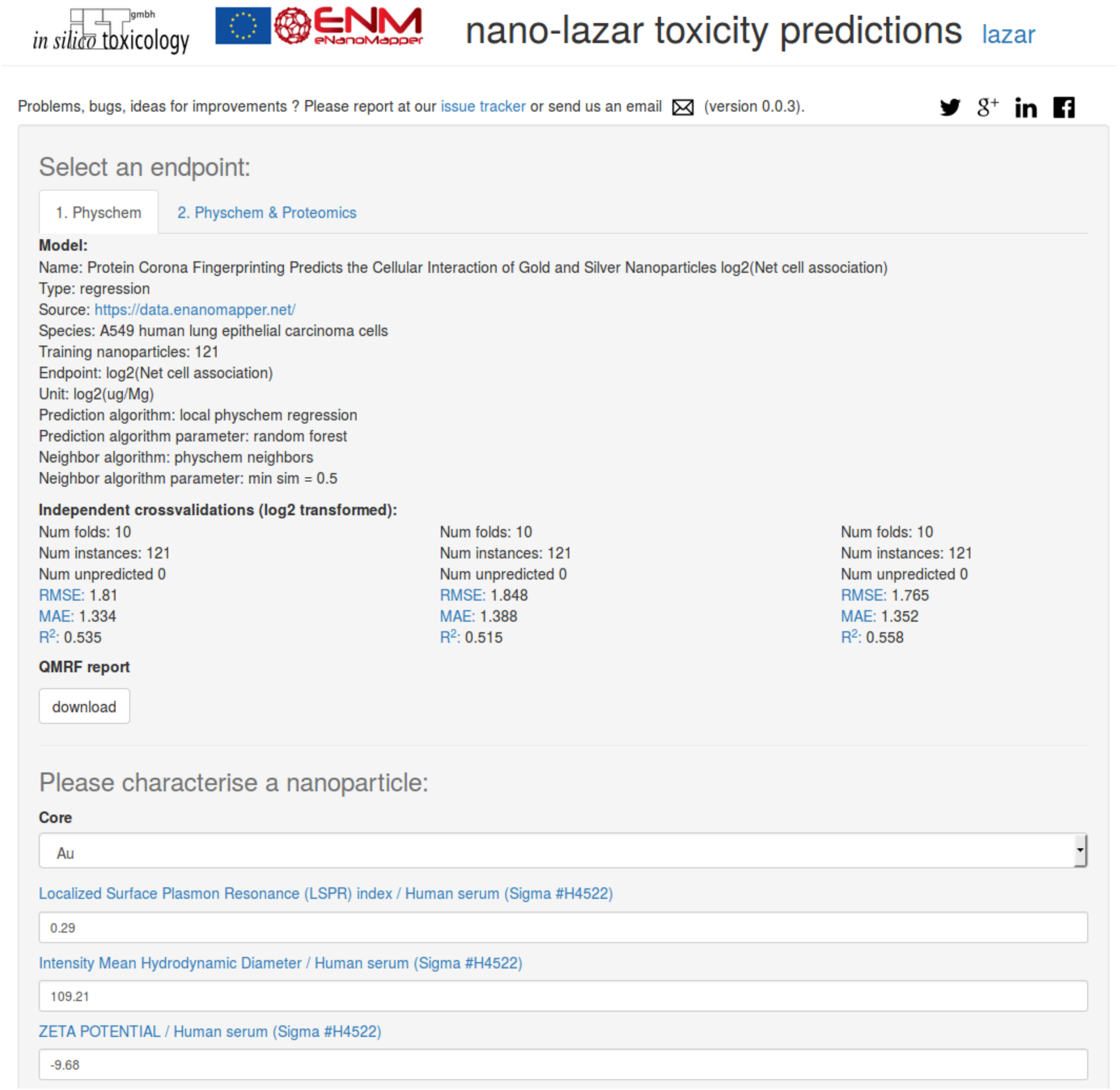
Chipster´s client software uses Java Web Start to install itself automatically and connects to computing servers to perform the actual analyses. Chipster server can be also installed (free of charge) in the EGI infrastructure for European researchers. EGI and ELIXIR are currently planning to set up also a ready-made Chipster server for end users.

- eNanoMapper resources: http://www.enanomapper.net/library/chipster-tool-analysis
- Chipster website: http://chipster.csc.fi/
- Contact developers: chipster@csc.fi

3.3.6. NANO-LAZAR

lazar is a modular framework for read across predictions of chemical toxicities. Within the eNanoMapper project lazar was extended with capabilities to handle nanomaterial data, interfaces to other eNanoMapper services (databases and ontologies) and a stable and user-friendly graphical interface for nanoparticle read-across predictions. See also Annex 6.1 for nano-lazar prediction results output.

- Access: https://nano-lazar.in-silico.ch
- Development: https://github.com/enanomapper/nano-lazar
- Feedback: https://github.com/enanomapper/nano-lazar/issues
- Technology: Web application, Library / Ruby
- License / Waiver: GPL3



nano-lazar toxicity predictions lazar

Problems, bugs, ideas for improvements ? Please report at our issue tracker or send us an email ✉ (version 0.0.3).

Select an endpoint:

Model:
Name: Protein Corona Fingerprinting Predicts the Cellular Interaction of Gold and Silver Nanoparticles log2(Net cell association)
Type: regression
Source: https://data.enanomapper.net/
Species: A549 human lung epithelial carcinoma cells
Training nanoparticles: 121
Endpoint: log2(Net cell association)
Unit: log2(ug/Mg)
Prediction algorithm: local physchem regression
Prediction algorithm parameter: random forest
Neighbor algorithm: physchem neighbors
Neighbor algorithm parameter: min sim = 0.5

Independent crossvalidations (log2 transformed):

Num folds: 10	Num folds: 10	Num folds: 10
Num instances: 121	Num instances: 121	Num instances: 121
Num unpredicted 0	Num unpredicted 0	Num unpredicted 0
RMSE: 1.81	RMSE: 1.848	RMSE: 1.765
MAE: 1.334	MAE: 1.388	MAE: 1.352
R^2: 0.535	R^2: 0.515	R^2: 0.558

QMRF report

Please characterise a nanoparticle:

Core
Au

Localized Surface Plasmon Resonance (LSPR) index / Human serum (Sigma #H4522)
0.29

Intensity Mean Hydrodynamic Diameter / Human serum (Sigma #H4522)
109.21

ZETA POTENTIAL / Human serum (Sigma #H4522)
-9.68

Figure 6. Nano lazar application

3.3.7. DESCRIPTORS AND MODELLING DEVELOPMENT TOOLS

3.3.7.1. GENE ONTOLOGY DESCRIPTOR (GO DESCRIPTOR) CALCULATION TOOL

An R package that calculates biological descriptors derived by clustering methodology of omics and Gene Ontology pathway data (*see* Deliverable 4.2).

- Technology: R
- License / Waiver: GPL3

3.3.7.2. NANO-LAZAR REST INTERFACE

The nano-lazar API Rest interface provides a Restful webservice and an interactive SWAGGER documentation to obtain OpenBabel, CDK and JoeLib descriptors for core and coating structures, QMRF reports and information on nano-lazar model, validation, dataset, feature and nanoparticles. Unit testing for nano-lazar rest is done with MiniTest library scripts.

- Access and API: https://enm.in-silico.ch/swagger
- Code: https://github.com/opentox/lazar-rest
- Feedback: https://github.com/opentox/lazar-rest/issues
- Testing: https://github.com/opentox/lazar-rest/tree/master/test
- Technology: Web application, Ruby
- License / Waiver: GPL3

3.3.7.3. CHEMISTRY DEVELOPMENT KIT DESCRIPTORS (STRUCTURAL DESCRIPTORS)

Nanomaterial descriptors implemented as part of the Chemistry Development Kit.

- Status: released
- Code: https://github.com/cdk/cdk
- Technology: Java
- License / Waiver: LPGL 2.1 or later

3.2.8. DOCUMENTATION ON MODELLING APPLICATIONS

~ Analysis and modelling

- Applications
    - ~ JaqPot Quattro
        - ~ API
        - ~ Feedback
        - ~ Source code
    - ~ Image Descriptor Calculation Web Tool
        - ~ Source code
    - ~ RRegrs Package
    - ~ ToxFlow Package
    - ~ Chipster for nanomaterial-based omics data analysis
    - ~ Nano-lazar toxicity predictions
        - ~ Feedback
    - Descriptors and development tools
        - ~ Gene Ontology descriptors
        - ~ Openbabel, Cdk and Joelib
        - ~ Chemistry Develop Kit descriptors (structural descriptors)
- User guidance
    - Tutorials
        - ~ Modelling Services
        - ~ JaqPot Quattro UI
        - ~ JaqPot Quattro API
        - ~ Image descriptor
        - ~ RRegrs package
        - ~ Chipster tool for analysis
        - ~ Pathway analysis
        - ~ How to make a pathway
        - ~ Statistics analysis
        - ~ AFFYQC web tool
    - Webinars
        - ~ Enriching protein corona
    - Publications
        - ~ Kohonen et al., 2014, Basic & Clinical Pharmacology & Toxicology
        - ~ Tsiliki et al., 2015, Journal of Cheminformatics
- ~ DoW
    - ~ Issue tracker
    - Reporting
        - ~ D4.1 Analysis and Modelling Specifications
        - ~ D4.2 Descriptor Calculation Algorithms and Methods
        - ~ D4.3 nQSAR Modelling infrastructure
        - D4.4 Mechanism-of-action Modelling Tools
        - D4.5 Design of experiments and inter-laboratory testing facilities
        - 4.6 Tools for generating QMRF and QPRF reports

3.4 SUPPORT APPLICATIONS

3.4.1. AUTHENTICATION & AUTHORIZATION INFRASTRUCTURE

During the whole project the infrastructure for authentication and authorization, user management and user registration was maintained and updated to newest security standards. User registration, authentication and authorisation was fully described in Deliverable D5.2.

3.4.2. SUMMIT APPLICATION

The web-based application was created as a collaborative and support tool for nano-safety related events, with possible extension to other areas.

- Version 1.0 (https://nanoehs.enanomapper.net/) was used to support the U.S-EU NanoEHS workhop and the scrimmage (Arlington, 2016)
- Version 2.0 (https://summit.enanomapper.net/) was used to support the EU-U.S. NanoEHS workshop and breakout session (Rheinfelden 2016)

The functionalities of the summit application facilitate the interactions before, during and after the event:

- Add and create content and resources (e.g. publications, guidance documents, tools)
- Add topics and questions to be discussed
- Add comments and answers related to the topics
- Gather information and capture the discussions
- Support the reporting

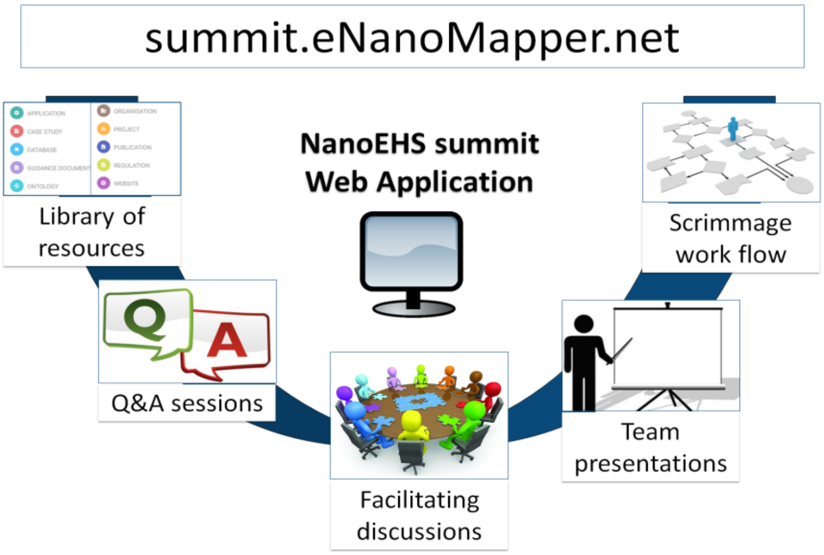


Figure 7. Summit application functionalities

3.4.3. DOCUMENTATION ON USER APPLICATIONS, DISSEMINATION AND TRAINING

~ User application development, integration & testing

- DoW
    - ~ Issue tracker
    - Reporting
        - ~ D5.1 Integrated Issue Management and Testing system
        - ~ D5.2 User registration, authentication and authorisation
        - ~ D5.3 User application for importing NanoWiki data
        - ~ D5.4 User application for importing NanoSafety Cluster data
        - ~ D5.5 User application for searching and downloading eNanoMapper data
        - D5.6 User application for conformance to reporting and curation standards
        - D5.7 Final report on User Applications
- Applications
    - Summit
        - ~ Version 1.0
        - ~ Version 2.0
- ~ GitHub

~ Dissemination & Training

- ~ Publications
- ~ Presentations
- ~ Tutorials
- ~ Webinars
- ~ Events
- DoW
    - ~ Issue tracker
    - Reporting
        - ~ D6.1 eNanoMapper Year 1 Dissemination Report
        - ~ D6.2 eNanoMapper Year 2 Dissemination Report
        - D6.3 eNanoMapper Tutorials
        - D6.4 eNanoMapper Community Development Report
        - D6.5 eNanoMapper Exploitation Report
        - D6.6 eNanoMapper Final Dissemination Report

4. PRIOTIZED CASE STUDIES

Use case	**Ontology**	**Database**	**Data Search**	**Protocol Service**	**JaqPot Quattro**	**RRegrs**	**Image Descriptor Tool**	**toxFlow**	**Chipster**	**Nano-lazar**	**Modelling descriptors**	**SPARQL endpoint**
02. Design protocol												
04. Upload protocol				**X**								
05. Upload dataset		**X**			**X**		**X**	**X**				
06. Design study					**X**							
07. Search protocol				**X**								
08. Search dataset		**X**	**X**		**X**							
08. Search ENM			**X**									**X**
10. Download dataset		**X**			**X**							
13. ISA-TAB creation *and/or export*		**X**										
14. Build QSAR					**X**	**X**	**X**			**X**	**X**	
17. Validate model					**X**	**X**		**X**		**X**		
20. Create template (xls etc)												
21. Map nanomaterial to biological pathway	**X**							**X**	**X**		**X**	
23. Search for all knowledge of nanomaterial X	**X**		**X**									
27. Find/ link related data sources		**X**						**X**			**X**	
29. Save my data in format X	**X**	**X**			**X**	**X**		**X**				
33. What URI to use for X												
34. Give me all names for ontology term A	**X**											
36. Harmonize Terminology	**X**											
38. Find ontological contradiction	**X**											
39. Comp aided ENM design					**X**						**X**	

44. Find producer of eNM C		X										
45. Find QC data of eNM C												
49. Map nano material found in DB 'A' to an entry in DB 'B'		X										
50. Find all DB's with data of Paper X			X									
52. Register and get access (public)		X			X							
53. Map existing used schemas to ontology IDs (e.g. OECD harmonized templates)	X							X			X	
54. Annotation of data with ontology IDs (e.g. experimental data in ISA-Tab)	X											
55. Experimental design					X							

5. CONCLUSION

A modular infrastructure was developed in order to achieve the main goal of the project in offering viable solutions to the nanosafety community regarding the data management (e.g. sustainable data storage) and data use (e.g. modelling). Therefore, eNanoMapper worked on developing of tools for data storage and analysis, information exchange and reporting of research (nano) data, but also towards creating a community framework for interdisciplinary collaboration in the area of nanosafety.

Therefore, the eNanoMapper team have: i) developed a nano ontology for the categorization and characterization of nanomaterials, ii) developed and implemented a modular infrastructure for data storage, sharing and searching, based on open standards and semantic web technologies, and iii) created and tested different computational modelling tools and services for nanomaterials safety assessment.

Overall, the application infrastructure developed within eNanoMapper is able to offer reliable data management solutions and facilitate the integration of experimental data into risk assessment by using different computational models.

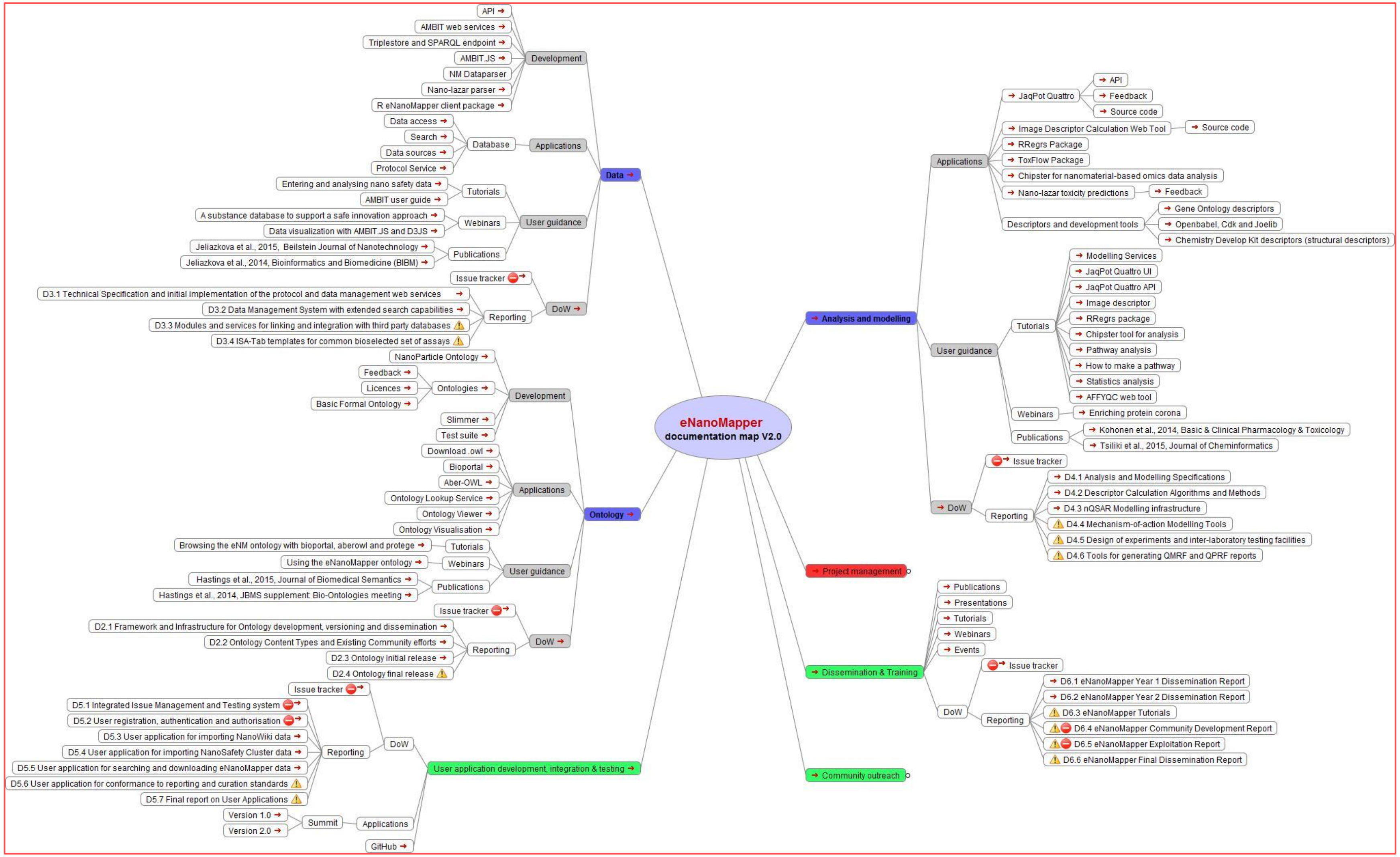
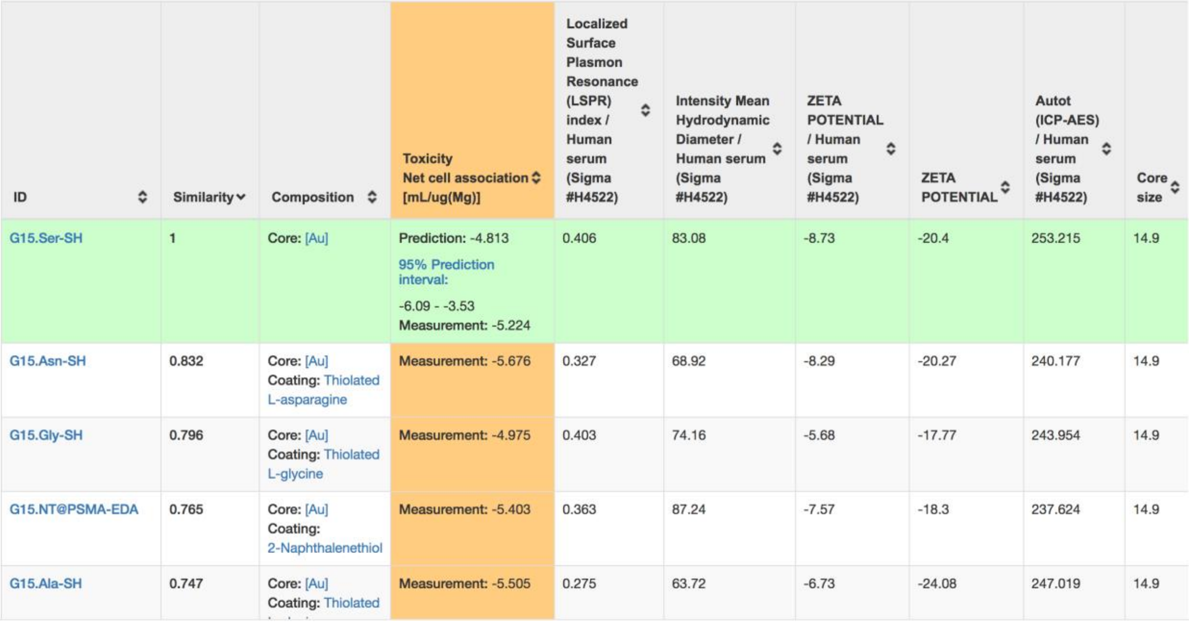


Figure 8. eNanoMapper documentation map

6. ANNEXES

6.1 NANO-LAZAR RESULTS OUTPUT

Nano-lazar presents predictions in tabular form with the query nanoparticle on top followed by similar nanoparticles for the local QSAR model. Each nanoparticle is characterised by its similarity to the query compound, endpoint measurements and the properties relevant for the particular endpoint. If the query nanoparticle is part of the training data (as in this example) all information from this substance is removed from the training set before making the prediction. In this case the measured value is displayed for comparison purposes.

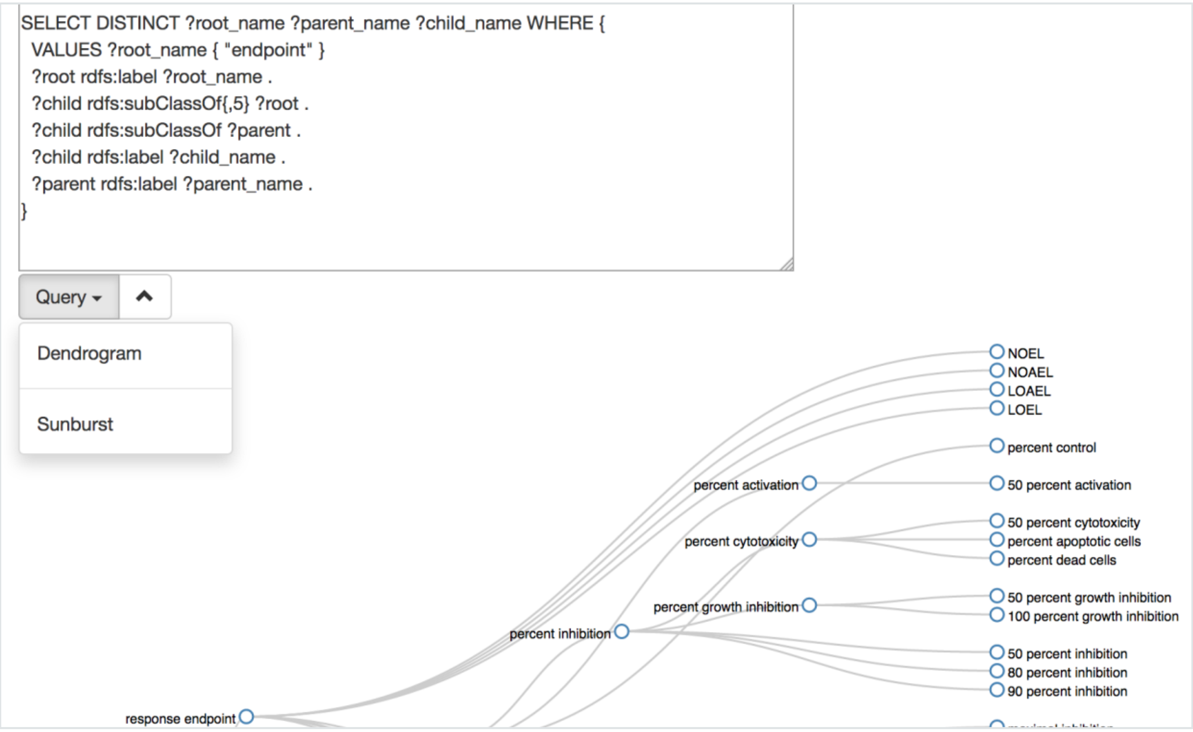


ID	Similarity	Composition	Toxicity Net cell association [mL/ug(Mg)]	Localized Surface Plasmon Resonance (LSPR) index / Human serum (Sigma #H4522)	Intensity Mean Hydrodynamic Diameter / Human serum (Sigma #H4522)	ZETA POTENTIAL / Human serum (Sigma #H4522)	ZETA POTENTIAL	Autot (ICP-AES) / Human serum (Sigma #H4522)	Core size
G15.Ser-SH	1	Core: [Au]	Prediction: -4.813 95% Prediction interval: -6.09 - -3.53 Measurement: -5.224	0.406	83.08	-8.73	-20.4	253.215	14.9
G15.Asn-SH	0.832	Core: [Au] Coating: Thiolated L-asparagine	Measurement: -5.676	0.327	68.92	-8.29	-20.27	240.177	14.9
G15.Gly-SH	0.796	Core: [Au] Coating: Thiolated L-glycine	Measurement: -4.975	0.403	74.16	-5.68	-17.77	243.954	14.9
G15.NT@PSMA-EDA	0.765	Core: [Au] Coating: 2-Naphthalenethiol	Measurement: -5.403	0.363	87.24	-7.57	-18.3	237.624	14.9
G15.Ala-SH	0.747	Core: [Au] Coating: Thiolated	Measurement: -5.505	0.275	63.72	-6.73	-24.08	247.019	14.9

Figure 9. Nano-lazar results output

6.2 ENM ONTOLOGY VIEWER GRAPHICAL AND FILE OUTPUT FORMATS

The eNM Ontology Viewer application is a three page, tab-structured interface. The “Viewer” page has a SPARQL interface contains a query template that collects all corresponding triples by the keyword “endpoint”. The user has the possibility to choose between two types of graphical representation of the query results. “Dendrogram” and “Sunburst” (Figure 10) .



```
SELECT DISTINCT ?root_name ?parent_name ?child_name WHERE {
 VALUES ?root_name { "endpoint" }
 ?root rdfs:label ?root_name .
 ?child rdfs:subClassOf{,5} ?root .
 ?child rdfs:subClassOf ?parent .
 ?child rdfs:label ?child_name .
 ?parent rdfs:label ?parent_name .
}
```

Figure 10. Select option for rendering results.

The output option “Dendogram” is a tree based vector graphic representing the relations of nodes in the triples of the query result (Figure 11). The request is handled in real-time by the client browser.

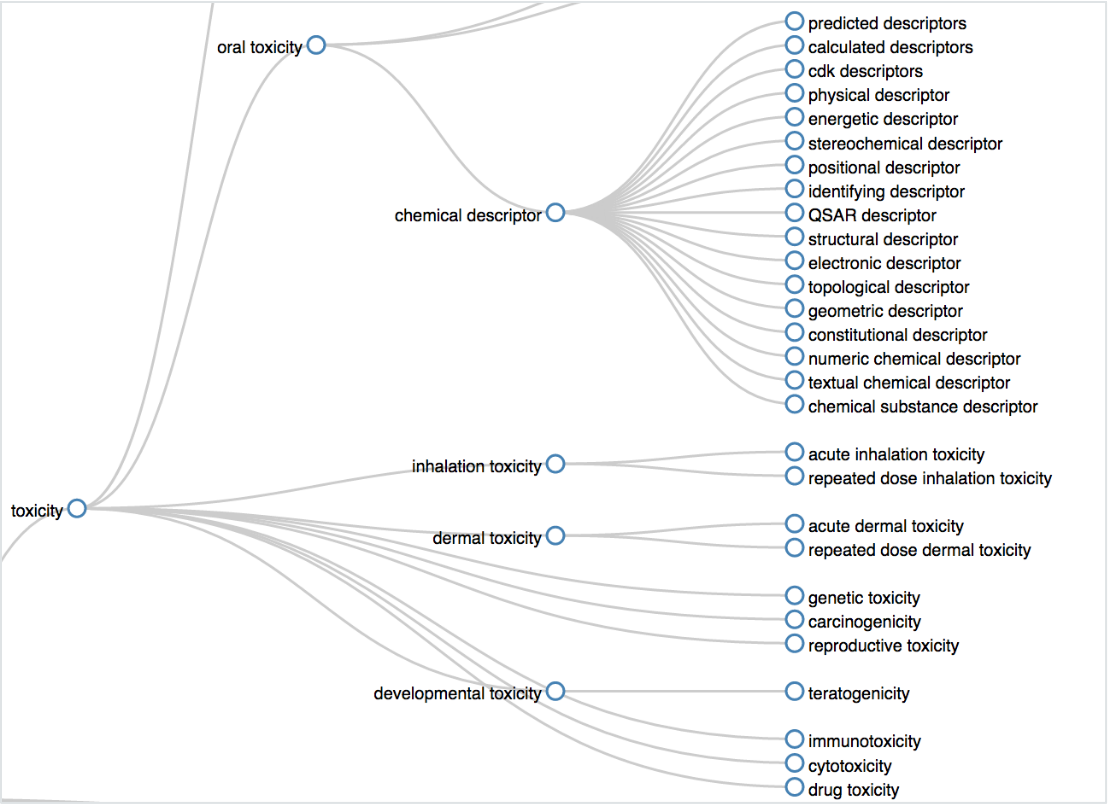


Figure 11. Dendrogram corresponding nodes

The “Sunburst” option generates an interactive pie chart vector graphic (Figure 12), that allows us to navigate through the levels of the result by clicking the colored fields in the graphic (Figure XY). Each color represents a binding of a node. By click on a colored field the graphic will zoom in so that this field gets the center of the graphic.

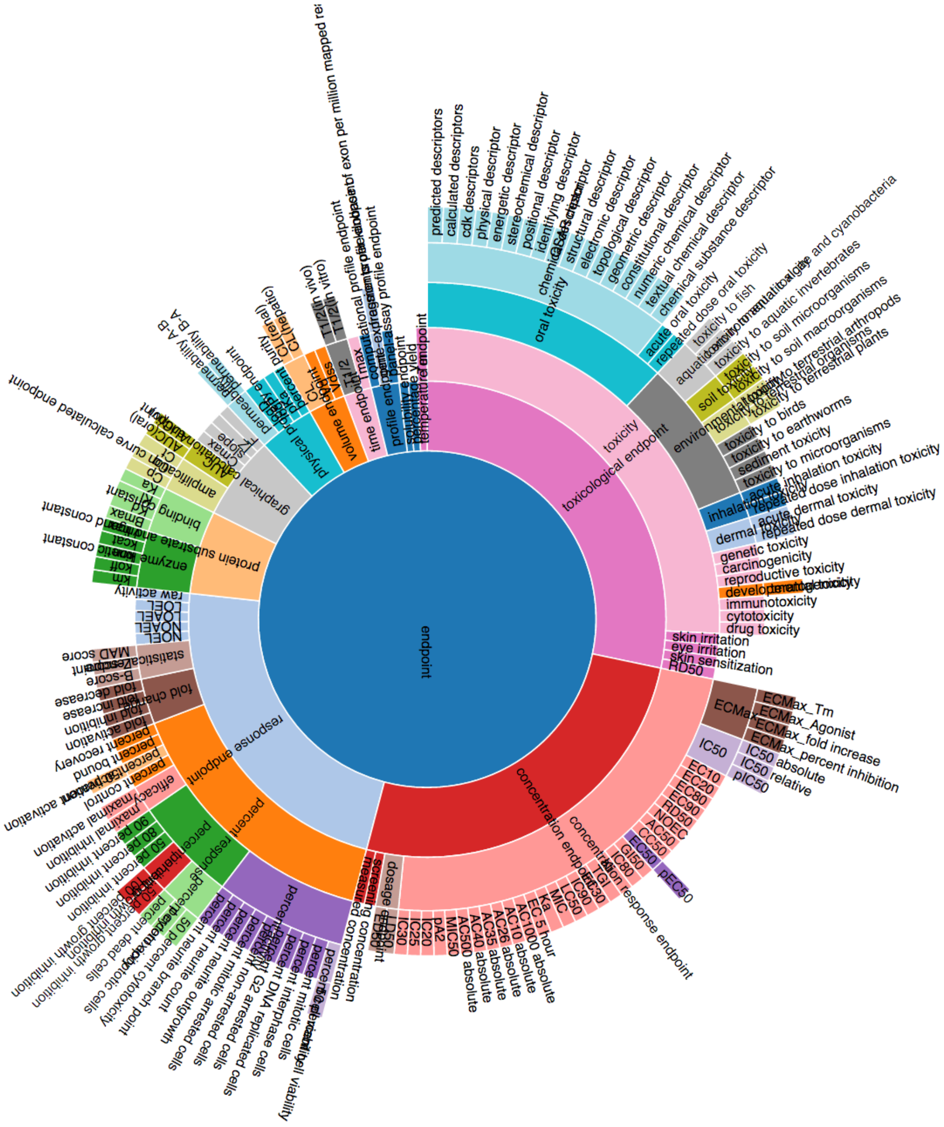


Figure 12. Sunburst interactive pie chart.

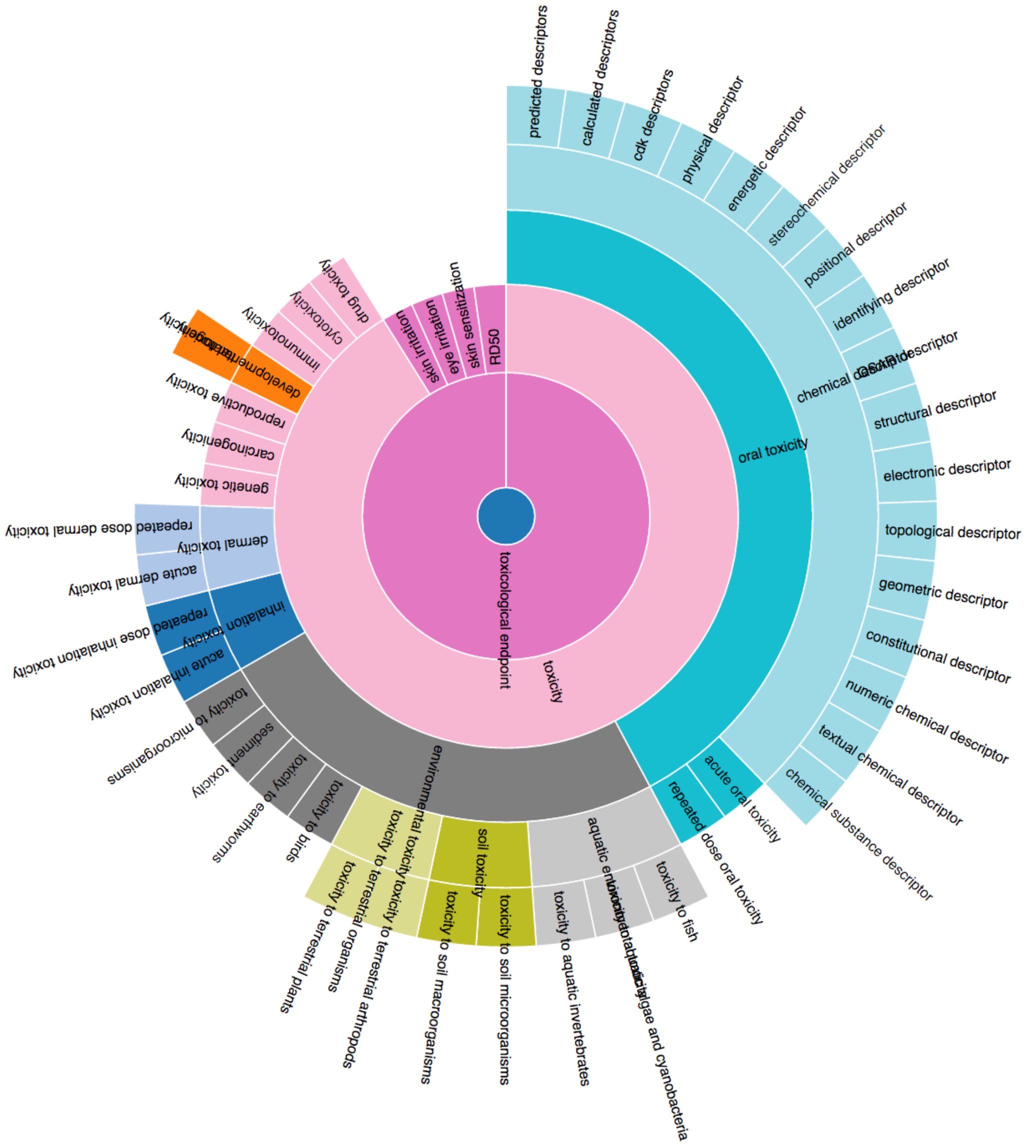
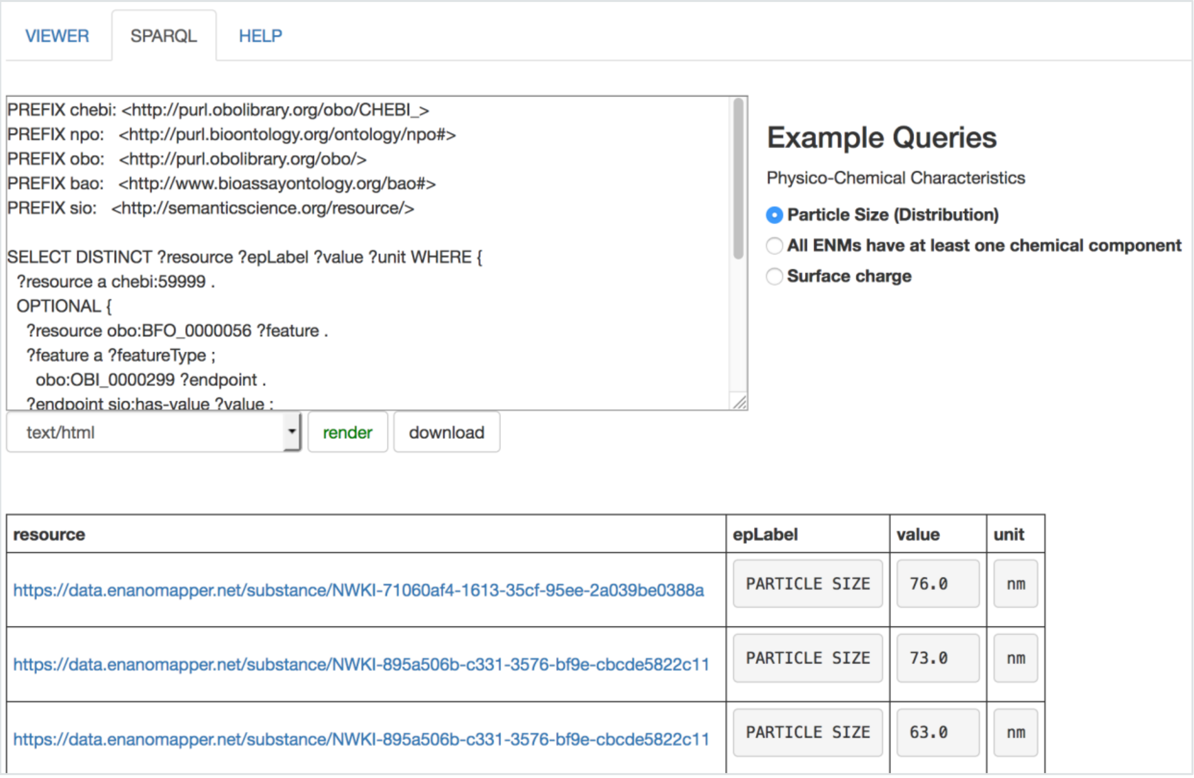


Figure 13. Click on colored fields brings up a deeper level of the pie chart

The “SPARQL” page contains a user interface (Figure 13) to execute SPARQL queries and download the result in various formats as a file. Furthermore, it contains some examples to select. “text/html” selected as output format also allows to render the result direct to the page. In this HTML table eNanoMapper ontology URIs link direct to the specific entry in the Bioportal service (http://bioportal.bioontology.org/ontologies/ENM). On Bioportal all visits to the eNanoMapper ontology get counted and visualized in a diagram.



```
PREFIX chebi: <http://purl.obolibrary.org/obo/CHEBI_>
PREFIX npo:   <http://purl.bioontology.org/ontology/npo#>
PREFIX obo:   <http://purl.obolibrary.org/obo/>
PREFIX bao:   <http://www.bioassayontology.org/bao#>
PREFIX sio:   <http://semanticscience.org/resource/>

SELECT DISTINCT ?resource ?epLabel ?value ?unit WHERE {
 ?resource a chebi:59999 .
 OPTIONAL {
  ?resource obo:BFO_0000056 ?feature .
  ?feature a ?featureType ;
   obo:OBI_0000299 ?endpoint .
  ?endpoint sio:has-value ?value ;
```

Example Queries

Physico-Chemical Characteristics

● Particle Size (Distribution)
○ All ENMs have at least one chemical component
○ Surface charge

resource	epLabel	value	unit
https://data.enanomapper.net/substance/NWKI-71060af4-1613-35cf-95ee-2a039be0388a	PARTICLE SIZE	76.0	nm
https://data.enanomapper.net/substance/NWKI-895a506b-c331-3576-bf9e-cbcde5822c11	PARTICLE SIZE	73.0	nm
https://data.enanomapper.net/substance/NWKI-895a506b-c331-3576-bf9e-cbcde5822c11	PARTICLE SIZE	63.0	nm

Figure 14. Graphical user interface to write SPARQL queries

A variety of output formats can be selected by a drop-down menu that is divided into SELECT and CONSTRUCT parts (Figure XY). The result clause (SELECT or CONSTRUCT) determines the associated output format that can be used.

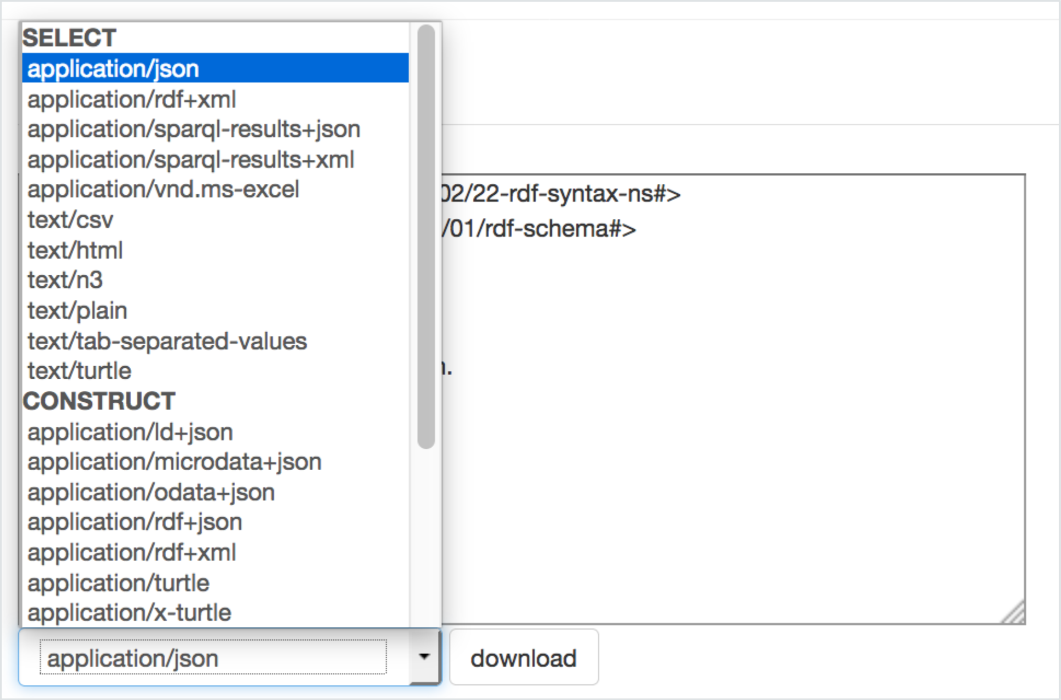


Figure 15. Select output format by SPARQL result clause