DELIVERABLE REPORT D1.4

D1.4 User Guidance

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:	IST
DATE:	31.10.2016
VERSION:	1.2

A Database and Ontology Framework for Nanomaterials Design and Safety Assessment

Call identifier	FP7-NMP-2013-SMALL-7
Document Type	Deliverable Report
WP/Task	WP1/T1.7
Document ID	eNanoMapper D1.4
Status	Final

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	Micha Rautenberg (IST), Lucian Farcal (DC), Philip Doganis (NTUA), Nina Jeliazkova (NTUA), Egon Willighagen (UM). Reviewed by Barry Hardy (DC).
Purpose of the Document	To report on User Guidance
Document History	1. Table of Contents, 25/08/2016 2. First draft, 08/11/2016 3. Final draft 31/10/2016 3. Complete version, 12/12/2016

TABLE OF CONTENTS

TABLE OF FIGURES

1. SUMMARY

The Provision of User Guidance (on data input, access to data, tool use, experimental design, model building) was developed from an early stage of development and was improved throughout the project with a final version that is made available in this deliverable, D1.4. User guidance and training was implemented through online tutorials, meetings, face-to-face workshops and conferences (organised by WP6). An eNanoMapper dictionary with the definitions of terms and abbreviations has been developed and deployed.

2. USER GUIDANCE

## 2.1 ONLINE RESOURCES AND TUTORIALS

2.1.1 OVERVIEW OF AVAILABLE ONLINE TUTORIALS

These online tutorials are fully described in Deliverable 6.3

- Modelling Services
  Create DataSet, Train a Model, Make a Prediction, Validate (split, cross or external)
- JAQPOT QUATTRO UI
  How to use the JaqPot Quattro User Interface to create and validate NanoQSAR models
- JAQPOT Quattro API
  How to use all available functionalities of the JaqPot Quattro web application
- Pathway analysis
  How to use the Pathway module of ArrayAnalysis.org for pathway analysis of microarray data
- How to make a pathway
  Creating biological pathways for WikiPathways using PathVisio
- Statistics analysis
  How to use the statistics module of ArrayAnalysis.org for statistics analysis of microarray data
- AffyQC web tool
  How to use the AffyQC web tool of ArrayAnalysis.org for quality control and pre-processing of Affymetrix microarray data
- Chipster Tool for Analysis
  How to Use Chipster for Bioinformatics Analysis of Nanomaterial-Based Omics Data
- Nano Safety Data
  Entering and Analysing Nano Safety Data
- Browsing the ontology
  Browsing the eNM ontology with BioPortal, AberOWL and Protégé
- Image descriptor Tutorial
  Image descriptor calculation web tool
- RRegrs: Package Tutorial

2.1.2 PRESENTATIONS AND SLIDES

NANOPARTICLES TOXICITY
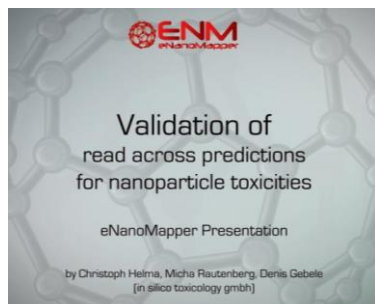
Validation of read across predictions for nanoparticle toxicities

Christoph Helma[1], Micha Rautenberg[1], Denis Gebele[1], [1]in silico toxicology gmbh, Basel, Switzerland

We validate lazar read-across models for nanoparticles, compare regression algorithms and compare nanoparticle descriptors. We provide an example for reproducible research.

Presenter: Christoph Helma (IST, CH)

ENABLING NANOSAFETY

eNanoMapper: enabling systems biology for nanosafety

E. Willighagen[1], F. Ehrhart[1], L. Rieswijk[1], B. Smeets[1], C. Evelo[1]

[1]Department of Bioinformatics - BiGCaT, Maastricht University

The eNanoMapper project develops computational solutions to support the development of new engineered nanomaterials within the European community. These solutions are based on a proposed common ontology, to unite the research done within the nanosafety domain.

This common language is based on reusing existing ontologies, like the NanoParticle Ontology (NPO), and developed in collaboration with other projects from the NanoSafety Cluster, but also by collaborating with other partners from inside and outside of Europe.

Another important component is the database approach: here, eNanoMapper selected the open and flexible OpenTox platform which has been extended to support nanomaterials: the user will find information including experimental data, physico-chemical and biological nanoparticle characterization.

By combining these two components and adopting open standards, such as ISATab-Nano, eNanoMapper is able to make user-oriented applications to support the day-to-day work of researchers in the community (and allow others to do so too).

NANOMATERIAL FRAMEWORK

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

Barry Hardy (Douglas Connect) at OpenTox Euro 2015 Meeting, UCD, Dublin, Ireland, 1 October 2015

This presentation provides an overview of the eNanoMapper project and framework which is developing a data management and analysis infrastructure together with ontologies supporting the safety assessment activities of the European nanomaterials research and development community. The project addresses the requirements of safety assessment of nanomaterials by providing databases, analysis tools and ontologies for risk assessment and linking them with existing resources in this area.

STORAGE AND QUERY

The eNanoMapper database for nanomaterial safety information: storage and query

Nina Jeliazkova[1], Nikolay Kochev[2], David Vorgrimmler[3], Janna Hastings[4], Vedrin Jeliazkov[1] Ideaconsult Ltd, Sofia, Bulgaria, [2]University of Plovdiv, Dep. of Analytical and Computer Chemistry, Plovdiv, Bulgaria, [3]in silico toxicology GmbH, Basel, Switzerland, [4]EMBL-EBI, Hinxton, United Kingdom

DATABASE AND ONTOLOGY

A Database and Ontology Framework for Nanomaterials Design and Safety Assessment ACS

Barry Hardy (Douglas Connect) at ACS Meeting, Boston, USA, 18 August 2015

This presentation provides an overview of the progress made in the initial 18 months of the project, including outlines of the initial development releases of the data platform and ontology.

ENANOMAPPER OVERVIEW

Nano Safety Cluster Meeting - Antalya, Turkey, 2014

Barry Hardy (Douglas Connect)

- Modular infrastructure for data storage, sharing and searching, based on open standards and semantic web technologies, minimum information standards and established security solutions;
- Development of ontologies for the categorisation and characterisation of nanomaterials in collaboration with other projects.
- Creation of new computational models in nanomaterials safety through the implementation of interfaces for toxicity modelling and prediction algorithms which may process all data made available through eNanoMapper (e.g. using algorithms available from the OpenTox FP7 project or statistical/data mining software).

Meta analysis of nano-bio interactions supporting "safe-by-design" ENMs development by pursuing a Linked Data approach which integrates data and metadata originating from diverse sources within nanoscience, chemistry, biology and toxicology.

Creation of tools for the exchange, quality assurance and reporting of research protocols and data for regulatory purposes. Creation of a community framework for interdisciplinary collaboration.

NANoREG Data Access

Presentations at NANoREG meetings (Nina Jeliazkova, IDEA)

- **Bilbao, Spain, July 2016**
- **The Hague, Netherlands, October 2016**

Enabling Systems Biology for Nanosafety

Presentation at ProSafe meeting (Egon Willighagen, UM)

- **Dessau, Germany, February 2016**

2.1.3 WEBINARS

ENRICHING PROTEIN CORONA

ENRICHING PROTEIN CORONA FINGERPRINTS USING GENE ONTOLOGY INFORMATION: AN INTEGRATION TECHNIQUE

The eNanoMapper project is working towards developing a modular and extensible infrastructure for data sharing, data analysis, and building computational toxicology models for ENMs. A number of services are now available for calculating ENM-specific descriptors, developing nanoQSAR models and automated workflows for model selection or validation.

Along these lines, we present a methodology to incorporate biological information with omics data and specifically proteomics data originating from protein corona fingerprinting, which has been reported to efficiently predict biological responses such as cellular uptake, signaling, and toxicity.

Our findings suggest that there is scope for further enhancement of protein corona data with biological information to allow for different protein weights according to their biological plausibility.

Presenter: GEORGIA TSILIKI (The National Technical University of Athens, Greece)

USING THE ONTOLOGY

USING THE ENANOMAPPER ONTOLOGY

Ontologies support a diverse range of applications including data integration, metadata enrichment (with synonyms and definitions) and standardisation.

The eNanoMapper project is developing an ontology driven by the principles of the OBO Foundry, most importantly of which being the re-use and integration of content from existing ontologies, which in the nano domain includes the Nanoparticle Ontology (NPO).

In this webinar, the structure and content of the eNanoMapper ontology was presented as well as a workflow for using the ontology in annotations and (if needed) requesting additional terms.

The webinar was demonstrated using the Protégé ontology editing and visualisation tool and the BioPortal online database of biomedical ontologies.

Presenter: JANNA HASTINGS (EMBL-EBI, UK)

DATA VISUALIZATION

NANOMATERIAL DATA VISUALIZATION WITH AMBIT.JS AND D3.JS
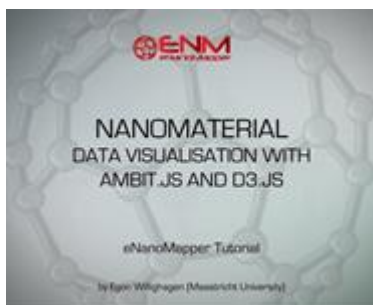
The EU FP7 eNanoMapper project is developing a data management and analysis infrastructure together with ontologies supporting the safety assessment activities of the European nanomaterials research and development community.

This webinar demonstrated how the eNanoMapper database Application Programming Interface (API) can be used to extract data to be visualized in graphs.

A short outline of the API was provided. On the client side, this API is used in this presentation by the ambit.js JavaScript library to select data which is then visualized using the d3.js library. Examples demonstrated how this technically works using the example eNanoMapper database instance.

Presenters: EGON WILLIGHAGEN (MAASTRICHT UNIVERSITY) and NINA JELIAZKOVA (IDEACONSULT LTD.)

ENANOMAPPER PROTOTYPE

A SUBSTANCE DATABASE TO SUPPORT A SAFE INNOVATION APPROACH

The webinar presents the first prototype released by the eNanoMapper project, providing support for upload and search for nanomaterials and experimental data through a REST web services API (http://enanomapper.github.io/API/) and a web browser interface. A number of opportunities and challenges exist in nanomaterials representation and integration of ENM information, originating from diverse systems, such as the concept of substances. We adopted the concept of substances, supporting ENMs as a special case; as well as experimental data, associated with substances. The substance resource supports assigning a nanomaterial type, chemical composition with relevant concentration and components' roles, as well as links to chemical structures. A file upload web form (or the underlying REST API call) allows importing files in several formats (OECD HT .i5z files or direct retrieval of information from IUCLDI5 servers); custom spreadsheet templates and custom formats, provided by partners (e.g. NanoWiki).

The eNanoMapper prototype database offers access to a variety of searches by combination of measurement endpoints (e.g. all ENMs with size between 50 and 60 nm and having genotoxicity data) and is tightly integrated with a chemical structure search. This allows searching for the component of a material using a chemical structure, and highlighting its function as a core, coating or functionalisation. The search can be used for many applications, one of which being NanoQSAR modeling. Here, the search interface is used to make a selection of data to be used for training and test data. Search results can then be downloaded in various formats and used in statistics and machine learning environments. Future work includes support for additional file formats, generated by the NanoSafety cluster, further annotation with ontologies and support for data from third party databases.

Presenters: NINA JELIAZKOVA (IDEACONSULT LTD.)

2.1.3.1 WEBINARS DESCRIBED IN OTHER DELIVERABLES

These webinars are already described in other deliverables:

- Nanotechnology and Safety
  Roland Grafström (VTT) 4th of March 2014
  described in D6.1
- The first eNanoMapper prototype: a substance database to support a safe innovation approach
  Nina Jeliazkova (IDEA) 29th of January 2015
  described in D6.1
- Enriching Protein Corona Fingerprints: an Integration Technique
  Georgia Tsiliki (NTUA) 29th Oct, 2015
  described in D6.2

- Using the eNanoMapper Ontology
  Janna Hastings, EMBL-EBI 25th June, 2015
  described in D6.2
- Nanomaterial Data Visualization with Ambit.js and D3.js
  Egon Willighagen, UM, and Nina Jeliazkova, IDEA 18th May, 2015
  described in D6.2

2.2 TOOLS USED IN FACE-TO-FACE WORKSHOPS AND CONFERENCES

ENANOMAPPER HANDS-ON WORKSHOP ON MODELLING RELATED TO NANO SAFETY, RHEINFELDEN, GERMANY, OCTOBER 25, 2016

The workshop had the format of several parallel sessions: following the structure of a Knowledge Café, where small groups work on specific problems under the guidance of a group leader within a specific time frame. Each exercise was repeated, so that the participants had the chance to navigate between tables and get trained in modelling tools. Applications developed within the eNanoMapper project were demonstrated. Participants had the possibility to use and test the tools, and also to discuss and give feedback to the developers. Details on this workshop including the materials used for the practical exercises can be found at this address: http://www.enanomapper.net/events/nano-modelling-workshop

- **Extracting knowledge from data using the JaqPot Modelling Tool,** Philip Doganis and Georgios Drakakis (National Technical University of Athens). The exercise offered hands-on work on the development of nanoQSAR models based on data available from the eNanoMapper database (https://data.enanomapper.net/), making use of the eNanoMapper computational infrastructure from National Technical University of Athens that extends the OpenTox API. The focus was on the use case of predicting cell association of metal oxide nanoparticles, based on publically available experimental data. Participants went through the workflow of constructing a model from a dataset drawn from the eNanoMapper database into the Jaqpot platform or added on the platform, getting predictions based on the model and evaluating its efficiency through model validation. Finally, users worked on the creation of predictive models using statistical and machine learning algorithms.
- **Nanomaterial read-across predictions with nano-lazar**, Christoph Helma (in silico toxicology). The exercise included a presentation of the lazar read-across framework, the adjustments of the application for nanoparticles, data requirements, comparison of algorithms and descriptors and exercises on gold and silver nanoparticles.

EU-US NANOEHS 2016 WORKSHOP, RHEINFELDEN, GERMANY, OCTOBER 24, 2016

- eNanoMapper, in cooperation with the BILAT USA 4.0 project, organized the workshop titled "Enabling a sustainable harmonised knowledge infrastructure supporting nano environmental and health safety assessment". The event served to support the ongoing US-EU dialogue in the area of Nano Environmental and Health Safety (NanoEHS) that has the goals to engage US and EU experts in active discussions, encourage joint programs of work and support communities of research. The workshop was attended by about 30 experts from Europe, U.S. and Asia.
- The workshop's main objective was to facilitate networking, knowledge sharing and idea development on the requirements and implementation of a sustainable knowledge infrastructure for NanoEHS. This infrastructure should support the needs required by different

stakeholders including researchers, industry, regulators, workers and consumers. The detailed information on the workshop can be found at this address: http://www.nanoehs-workshop.eu/

- The workshop was chaired by Dr. Barry Hardy (Coordinator of eNanoMapper) and Dr. Lisa Friedersdorf (National Nanotechnology Coordination Office, USA) and included presentations and discussions under the following sessions:
  - Information Needs of Stakeholders & Applications
  - US-EU perspectives, initiatives and progress
  - Knowledge Infrastructure Solutions
- The presentations were followed by two interactive Knowledge Café sessions:
  - **Read-across and data completeness**, session led by Dr. Frederick Klaessig (Pennsylvania Bio Nano Systems, USA). The objective was to broaden awareness of data completeness while fostering best practices within the US-EU communities. The participants discussed the data requirements and gaps for read-across applicability with nanomaterials, the concepts of data completeness from the recent nanoWG series of papers and the use of current templates to guide data collection. Zeta potential was used as a case study, with a focus on data taken for the NM-series and related materials. The participants surveyed several aspects and then discussed the relative merits of isoelectric point versus zeta potential for identifying coatings on particles, current templates and requirements on pH values for zeta potential measurements, solubility, data quality scores and the need for minimum ontology for coated particles that drives the testing and metadata. Overall, the importance of guidance on the purpose, tools and interpretations of zeta potential and other p-chem parameters was explored.
  - **Nano Ontology Evaluation**, session led by Egon Willighagen (Maastricht University, The Netherlands). eNanoMapper and other initiatives have developed a substantial ontology to be used in nanotechnology and safety assessment. This ontology can be used for harmonization purposes, ensuring a common description and reporting format. Based on the case study interaction and discussions, the participants evaluated how the ontology is supporting the scientific and regulatory needs, what the gaps are and identified further actions. Guided by a detailed tutorial, the participants were able to browse the eNanoMapper ontology, use the ontology in searching on http://search.data.enanomapper.net, and finally evaluate the appropriateness of the eNanoMapper ontology by completing an evaluation survey.

- The participants were able to use an interactive Knowledge Integrating web tool "Summit" developed by Douglas Connect (https://summit.enanomapper.net/), in order to access specific resources, and actively contribute to the discussions by adding new content, comments and answers related to the session topics.

ENANOMAPPER HANDS-ON WORKSHOP, ATHENS, GREECE, SEPTEMBER 29-30, 2016

eNanoMapper partners from NTUA, Misvik Biology and IST described specific cases and practical examples in which the group used tools to work through modelling, analysis and assessment exercises, and discussed the results. Details on this workshop including the materials used for the practical exercises can be found at this address: http://www.enanomapper.net/events/hands-on-workshop-athens

- **Extracting knowledge from data using the JaqPot Modelling Tool**
  Philip Doganis -National Technical University of Athens
- **Omics Data Analysis for Nanotoxicology**
  Penny Nymark - Misvik Biology Oy
- **Using KNIME for modelling toxicity in nanoparticles**
  Georgios Drakakis - National Technical University of Athens
- **Analyse nanotoxicity data using R**
  Georgia Tsiliki - National Technical University of Athens
- **Filling knowledge gaps in Nanotoxicology with Read-across Predictions, Practice & Requirements**
  Christoph Helma - in-silico toxicology gmbh

2ND NANOSAFETY FORUM FOR YOUNG SCIENTISTS, SEPTEMBER 15-16, 2016

During the young scientists meeting in Visby, an interactive exercise was conducted with participants in the meeting, where we invited participants to contribute data. This data was curated and annotated with ontology terms and available in an instance of the eNanoMapper database software. Full details are provided in D3.3.

PRACTICAL HANDS-ON WORKSHOP ON NANOSAFETY ASSESSMENT, SATELLITE SESSION AT THE NANOKOREA 2016, KINTEX, KOREA, JULY 13, 2016

- **Practical Session 1 - Data Management**, Nina Jeliazkova, IdeaConsult Ltd
  The goal of this workshop was to make the participants familiar with the eNanoMapper solutions for data management and data access. We 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.
- **Practical Session 2 - Data Modelling**, Barry Hardy, Douglas Connect GmbH
  During the session data requirements for nanotoxicity read across predictions and the coverage of currently available datasets was discussed. Participants were able to evaluate a prototype nanoparticle read across application and provide feedback for future developments. We explained how to develop nanoQSAR models based on data available from the data.enanomapper.net server and using the eNanoMapper computational infrastructure and demonstrated how a modeller can use diverse sets of nanoparticle descriptors, including descriptors derived from protein-corona information.

NANOMILE - ENANOMAPPER MEETING, EDINBURGH, APRIL 5-7, 2016

In this meeting the eNM partners from UM met with the NanoMILE team (people from University of Birmingham, National Hellenic Research Foundation (Greece) and the Karlsruhe Institute for Technology (Germany)) to discuss the eNM database and ontology and how these would fit into the NanoMILE project. Finally, suggestions and questions for improving both the ontology as well as the database were raised. Also, three presentations were given:

- **The NanoMILE project: Database and ontology synergies,** by Iseult Lynch from the University of Birmingham.

- **The eNanoMapper Ontology**, by Linda Rieswijk from the Maastricht University.
- **The eNanoMapper Database**, by Friederike Ehrhart from the Maastricht University.

NECID-ᴇNᴀɴoMᴀᴘᴘᴇʀ Oɴᴛoʟoɢʏ Woʀᴋsʜoᴘ, 3 Mᴀʀᴄʜ 2016, Lᴇɪᴅᴇɴ, NL

A workshop dedicated to the NANoREG project was held in December 2015 at the TNO buildings in Leiden, the Netherlands. The morning session consisted of tutorials detailing various of the solutions developed in eNanoMapper. The participants worked in the afternoon session on putting those solutions to use for the NANoREG project. Below is a list of talks given:

- **A general introduction to semantic web and ontology,** Egon Willighagen (University of Maastricht, eNanoMapper)
  - Discussing the benefits of using ontologies in modern databases.
- **The eNanoMapper ontology**, Gareth Owen (EMBL-EBI, eNanoMapper)
  - An explanation of the specific concept, intention and function of the eNanoMapper ontology, adoption of 3rd party content.
- **ISA-TAB/ISA-TAB-Nano and the new ISA-JSON specification**, Nikolay Kochev (IdeaConsult Ltd.)
  - Discussing the output options, Java classes, mapping to the eNanoMapper data model.
- **How to browse ontologies on online repositories like BioPortal or AberOwl,** Linda Rieswijk and Freddie Ehrhart (Maastricht University)
  - How to download ontologies and look up identifiers, how to download and browse ontologies with the open-source Protegé software, etc.
- **Configurable parser for Excel templates**, Nikolay Kochev (IdeaConsult Ltd.)
  - Creating JSON configuration for NanoSafety cluster templates.

ᴇNᴀɴoMᴀᴘᴘᴇʀ Hᴀɴᴅs-oɴ Woʀᴋsʜoᴘ oɴ Nᴀɴo Sᴀғᴇᴛʏ Assᴇssᴍᴇɴᴛ, Bᴀsᴇʟ, Swɪᴛᴢᴇʀʟᴀɴᴅ, Fᴇʙʀᴜᴀʀʏ 10, 2016

- **Nano Safety Science & Assessment**, Barry Hardy (Douglas Connect / eNanoMapper) and Vladimir Lobaskin (UCD)
- **Entering and Analysing Nano Safety Data**, Egon Willighagen (University of Maastricht) and Nina Jeliazkova (Ideaconsult / eNanoMapper)
  Workshop participants worked through the preparation of datasets to be used for modelling and risk assessment. The combination of data from multiple sources using common harmonised ontology and templates have been carried out by the group.
- **Data Requirements for Nanotoxicity Read-across Predictions**, Christoph Helma (in silico toxicology / eNanoMapper)
  Data requirements for nanotoxicity read across predictions and the coverage of currently available datasets were discussed. Participants have been able to evaluate a prototype nanoparticle read across application and provide feedback for future developments. Nano-Lazar Toxicity Predictions https://nano-lazar.in-silico.ch/predict
- **Bioinformatics Analysis of Nanotechnology-based Omics Data**, Pekka Kohonen and Penny Nymark (Misvik Biology Oy / eNanoMapper)
  Bioinformatics analysis of an omics dataset generated on a nanoparticle category were carried

out and the dose-dependence and mechanistic interpretations discussed.
[Tutorial](#)[1] | [Data](#)[2]

- **Developing Nano-QSAR predictive toxicology models**, Philip Doganis and Haralambos Sarimveis (NTUA / eNanoMapper)
  Participants worked on developing nanoQSAR models based on data available from the [data.enanomapper.net](#)[3] server and using the eNanoMapper computational infrastructure. We demonstrated how a modeler can use diverse sets of nanoparticle descriptors, including descriptors derived from protein-corona information and produce mathematical models of nanomaterial behaviour.
  [PDF](#)[4], [11_00 Developing Nano-QSAR predictive toxicology models.xml](#)[5]

- **Modelling Nano Exposure**, Claus Svendsen (UK Natural Environment Research Council / GUIDEnano and NanoFASE) and Wouter Fransman (TNO / GUIDEnano)
  The participants worked through the modelling of two exposure scenarios a) human (occupational and consumer) exposure to nanoparticles across different scenarios in the lifecycle of a product, and b) life cycle analysis of nanoparticles in the environment.
  [NECID Guidance](#)[6], [Stoffenmanager Nano](#)[7], [Case 1](#)[8], [Case 2](#)[9], [Excercise Soil](#)[10], [Excercise River](#)[11]

NANoREG-eNanoMapper Ontology Workshop, 2 December 2015, Leiden, NL

A workshop dedicated to the NANoREG project was held in December 2015 at the TNO buildings in Leiden, The Netherlands. The morning session consisted of tutorials detailing various of the solutions developed in eNanoMapper. The participants worked in the afternoon session on putting those solutions to use for the NANoREG project. Below is a list of talks given:

- **A general introduction to semantic web and ontology,** Egon Willighagen (University of Maastricht, eNanoMapper)
  - Discussing the benefits of using ontologies in modern databases.
- **The eNanoMapper ontology**, Gareth Owen (EMBL-EBI, eNanoMapper)
  - An explanation of the specific concept, intention and function of the eNanoMapper ontology, adoption of 3rd party content.
- **ISA-TAB/ISA-TAB-Nano and the new ISA-JSON specification**, Nikolay Kochev (IdeaConsult Ltd.)
  - Discussing the output options, Java classes, mapping to the eNanoMapper data model.

---

[1] http://enanomapper.net/docs/Chipster_tutorial_Workshop_Basel_100216.pdf

[2] http://enanomapper.net/docs/Chipster_session_Workshop_Basel_100216_hands_on.zip

[3] https://data.enanomapper.net

[4] https://drive.google.com/file/d/0BwmQMY3mRXTUQldCVUNiSHlHZ28/view

[5] https://drive.google.com/file/d/0BwmQMY3mRXTUNGlkMjdNQUZERVE/view

[6] http://enanomapper.net/docs/Guidance-NECID_version%201.0.pdf

[7] https://nano.stoffenmanager.nl/

[8] http://enanomapper.net/docs/Case1.pdf

[9] http://enanomapper.net/docs/Case2.pdf

[10] http://enanomapper.net/docs/NanoFASE_Class-exercise_soil-NSCWSBaselFeb2016.pdf

[11] http://enanomapper.net/docs/NanoFASE_Class-exercise_river-NSCWSBaselFeb2016.pdf

- **How to browse ontologies on online repositories like BioPortal or AberOwl,** Linda Rieswijk and Freddie Ehrhart (Maastricht University)
  - How to download ontologies and look up identifiers, how to download and browse ontologies with the open-source Protegé software, etc.
- **Configurable parser for Excel templates**, Nikolay Kochev (IdeaConsult Ltd.)
  - Creating JSON configuration for NanoSafety cluster templates.

OPENTOX EURO 2014, 22-24 SEPTEMBER 2014, ATHENS, GREECE

A Workshop was included in the OpenTox 2014 Conference (Athens, Greece, September 2014), event co-sponsored by eNanoMapper and ToxBank Projects. The workshop sessions were led by different eNanoMapper partners:

- · Chemical substances, nanomaterials and endpoint data in AMBIT (by **IDEA**),
- · Development of predictive NanoQSAR models using OpenTox infrastructure and the R language (by **NTUA**),
- · Open science pathway analysis (by **UM**) and
- · Carrying out a meta-analysis across multiple heterogeneous sources of evidence (by **DC**).

More information is available here: http://www.opentox.net/events/opentox-euro-2014

Furthermore, during the conference a Knowledge Cafe session took place, where in two sessions participants got the chance to discuss eNanoMapper tools with project partners:

Knowledge Cafe Session	Moderator
Data resources and management	Nina Jeliazkova
OpenTox APIs	Egon Willighagen

More information is available here: http://www.opentox.net/opentox-euro-sessions/workshop

2.3 OTHER TOOLS AND RESOURCES

2.3.1 ENM GITHUB REPOSITORIES

eNanoMapper supports developers in the area with a broad range of documentation, blog entries and tutorials on the libraries and source code output of the project. Most of the eNanoMapper source code is stored and documented at Github. Github is a 'web-based repository hosting system with distributed version control and source code management functionality'. The eNanoMapper project directory on Github at https://github.com/enanomapper has over 26 repositories with 23 active development members grouped in 7 developer teams. Each repository has an associated issue tracker with an overall sum of 230 issues, created by developers and users.

For the archival and versioning of, and collecting reported issues on eNanoMapper tutorials, there is a separate tutorial repository at https://github.com/enanomapper/tutorials.

2.3.2 TECHNICAL INFORMATION FOR DEVELOPERS

eNanoMapper provides a wide variety of information for software developers. Diverse blog entries, code documentations and publicly available reports on technical specification give a solid knowledge base for developers and system-integrators.

- API documentation in the Swagger specification provided in Swagger UI interfaces:
  - eNanoMapper database API http://enanomapper.github.io/API/
  - JQ Read-across API http://test.jaqpot.org:8080/jaqpot/swagger/
  - lazar and nano-lazar Rest API https://enm.in-silico.ch/swagger/
- OpenTox Github pages http://opentox.github.io/archive with information on implementation and system adjustment.
- RubyDoc.info with information on the lazar framework for the lazar, nano-lazar, lazar-gui, lazar-rest and QSAR-report libraries. All Ruby based code in the lazar framework is consistently documented with the YARD[12] documentation generation tool.
- Documentation on JAQPOT QUATTRO web application is described In Deliverable D4.3.

2.3.3 ENM DICTIONARY

The dictionary includes a collection of terms and abbreviations often used in the context of the eNanoMapper project, but also terms that are generally used in areas like nanosafety, data management or computational modelling. The eNanoMapper Dictionary in Version 1.0 can be downloaded in the online library[13]. The dictionary can also be found in Annex 5.1.

---

[12] YARD documentation tool http://yardoc.org/

[13] eNanoMapper Dictionary http://enanomapper.net/library/enm-dictionary

3. ANNEXES

3.1 eNM Dictionary

Term or abbreviation	Explanation
AMBIT.Js	*A Java Script Library for interaction with AMBIT instances;*
AMBIT	*AMBIT (http://ambit.sf.net) is an open source web application for cheminformatics data management, modelling and read across. It offers graphical user interface and REST web services, including implementation of OpenTox API and eNanoMapper API extensions. The eNanoMapper database http://data.enanomapper.net runs a customized version of AMBIT.*
API	**A**pplication **P**rogramming **I**nterface *A way computer programs talk to one another. Can be understood in terms of how a programmer sends instructions between programs. The API specifies how software components should interact. A good API makes it easier to develop a program by providing all the building blocks. A programmer then puts the blocks together.*
ArrayAnalysis.org	*Offers user-friendly solutions for gene expression data analysis, from raw data to biological pathways. It contains modules of three types that can be launched individually or successively as an integrated workflow; QC and pre-processing, statistical analysis and pathway analysis.*
ArrayExpress	*A database with data from high-throughput functional genomics experiments stored in a standardized format, providing these data for reuse to the research community. It also includes data on nanomaterials.* *https://www.ebi.ac.uk/arrayexpress/*
Bioclipse	*A Java-based, open source advanced workbench for chem- and bioinformatics. It provides 2D-editing and 3D-visualization of molecules, proteins and sequences, calculation of chemical properties, QSAR and much more; all fully integrated into a user-friendly desktop application. http://www.bioclipse.net/*
caNanoLab	**ca**ncer **Nano**technology **Lab**oratory *is a data sharing portal providing support for the annotation of nanomaterials with characterizations resulting from physico-chemical,* in vitro *and* in vivo *assays and the sharing of these characterizations and associated nanotechnology protocols in a secure fashion. https://cananolab.nci.nih.gov/caNanoLab/*
ChEMBL	*A manually curated chemical database of bioactive molecules with drug-like properties developed at the European Molecular Biology Laboratory. It also contains data on nanomaterials. https://www.ebi.ac.uk/chembl/*
Chipster	*An open source user-friendly analysis software for omics and high-throughput data. It contains over 350 analysis tools for next generation sequencing (NGS), microarray, proteomics and sequence data. Users can save and share automatic analysis workflows, and visualize data interactively.*
CTD	**C**omparative **T**oxicogenomics **D**atabase *is a manually curated robust, publicly available database providing information about chemical–gene/protein*

	interactions, chemical–disease and gene–disease relationships, integrated with functional and pathway data to aid in development of hypotheses about the mechanisms underlying environmentally influenced diseases. It also contains data on nanomaterials. http://ctdbase.org/
eNM	*eNanoMapper – A Database and Ontology Framework for Nanomaterials Design and Safety Assessment (EU Project)*
GEO	*A database with data from high-throughput functional genomics experiments stored in a standardized format, providing these data for reuse to the research community. It also includes data on nanomaterials. http://www.ncbi.nlm.nih.gov/geo/*
ISA-TAB	The **I**nvestigation / **S**tudy / **A**ssay **tab** - delimited *format is a general purpose framework with which to collect and communicate complex metadata (i.e. sample characteristics, technologies used, type of measurements made) from experiments employing a combination of technologies. In particular, ISA - TAB has been developed for —but not limited to — experiments using genomics, transcriptomics, proteomics or metabol/nomics techniques.* *For example, consider an investigation into the effect of a compound that induces liver damage, which looks at changes in (i) the metabolite profile of urine and (ii) gene expression in the liver (by mass spectrometry and microarray technologies, respectively). The general motivation for this work is the fulfillment of the needs of two group. http://www.isa-tools.org/*
ISO 13528:2015	*Statistical methods for use in proficiency testing by interlaboratory comparison; from International Standard Organization (ISO)*
IUCLID5	**I**nternational **U**niform **C**hemical **I**nformation **D**atabase version **5** *plays a central role in the IT environments of all organisations that have to cope with data submission requirements of REACH and other programs (OECD HPV, EU Biocides and others). Industry stakeholders, EU Member States, the European CHemicals Agency (ECHA) , and any other interested party obtain the IUCLID installation kit from the web site; once installed, the local IUCLIDs are the essential tool to capture & store, submit, and exchange data on chemical substances stored according the format of the OECD Harmonised Templates http://iuclid.eu/*
Jaqpot Quattro	*An eNanoMapper developed web application that currently supports data preprocessing, as well as statistical, data mining and machine learning algorithms and methods for defining the Domain of Applicability of a predictive nQSAR model. It is an extension of the Jaqpot web application, which was originally developed during the OpenTox project.*
JSON	**Java**S**cript **O**bject **N**otation *is a lightweight data-interchange format. It is easy for humans to read and write. It is easy for machines to parse and generate. JSON is a text format that is completely language independent but uses conventions that are familiar to programmers of the C-family of languages, including C, C++, C#, Java, JavaScript, Perl, Python, and many others. It is based on a subset of the JavaScript Programming Language, Standard ECMA-262 3rd Edition - December 1999. http://json.org/*
MOPAC	**M**olecular **O**rbital **PAC**kage *is a semiempirical quantum chemistry program for generating QSAR descriptors.*
MySQL	**M**y-**S**tructural **Q**uery **L**anguage *is an open-source relational database management system, enabling delivery of reliable, high-performance and*

	scalable Web-based and embedded database applications. "My" is the name of daughter to one of the founders. https://www.mysql.com/
Nano-lazar	**Nano-Laz**y structure-**a**ctivity **r**elationships *is an eNanoMapper developed tool for read across toxicity prediction (*https://nano-lazar.in-silico.ch/predict*).*
NanoWiki	*Originally developed as an internal knowledgebase of the toxicity of, primarily, metal oxides at the Karolinska Institutet and Maastricht University. Uses* https://semantic-mediawiki.org/ *for data input. The data is exported as RDF dump , available at* http://dx.doi.org/10.6084/m9.figshare.1330208. *The NanoWiki RDF dump is converted with a custom parser and imported into the eNanoMapper database at* http://data.enanomapper.net
OECD HT	The **OECD H**armonized **T**emplates (OHTs) *are structured (XML) data formats for reporting safety-related studies on chemical substances. The OHTs* http://www.oecd.org/ehs/templates/ *) and the supporting IT tool (IUCLID5,* http://iuclid.eu *) are used in a regulatory context, for preparation of substance dossiers for REACH and for other regulatory frameworks operating in Europe; as well as by the JRC NanoHub database*
Ontology	*A relational controlled vocabulary; deals with questions concerning what entities exist or may be said to exist, and how such entities may be grouped, related within a hierarchy, and subdivided according to similarities and differences.*
OpenTox	*A predictive toxicology framework with a unified access to toxicological data, (Q)SAR models and supporting information developed under the grant agreement FP7-HEALTH-2007-A- 200787*
Open Source	**Open-source** software (OSS) *is computer software with its source code made available with a license in which the copyright holder provides the rights to study, change, and distribute the software to anyone and for any purpose. Open-source software may be developed in a collaborative public manner.*
PathVisio	*An open-source and free pathway analysis and pathway drawing software. Allows drawing, editing and analysis of biological pathways.* http://www.pathvisio.org/
PROAST	*An R-based software package that has been developed by RIVM for the statistical analysis of dose-response data. It can be used for 1) dose-response modelling, 2) deriving a Bench Mark Dose in human risk assessment, and 3) deriving an effect concentration in ecotoxicological risk assessment.*
PubChem	The **PubChem** Compound Database *contains validated chemical depiction information provided to describe substances in PubChem Substance. Structures stored within PubChem Compounds are pre-clustered and cross-referenced by identity and similarity groups. It also contains data on nanomaterials.* https://pubchem.ncbi.nlm.nih.gov/
RDF	The **R**esource **D**escription **F**ramework *is one of the core common standards and data exchange formats of the Semantic Web. RDF it is based upon the idea of making statements about resources (in particular web resources) in the form of subject–predicate–object expressions. These expressions are known as triples in RDF terminology.* https://www.w3.org/RDF/ *Example: "The sky has the color blue" in RDF is as the triple: a subject denoting "the sky", a predicate denoting "has", and an object denoting "the color blue".*

REST	**Re**presentational **s**tate **t**ransfer *is an abstraction of the architecture of the World Wide Web; more precisely, REST is an architectural style consisting of a coordinated set of architectural constraints applied to components, connectors, and data elements, within a distributed hypermedia system. REST ignores the details of component implementation and protocol syntax in order to focus on the roles of components, the constraints upon their interaction with other components, and their interpretation of significant data elements. OpenTox web services are based on the REST architecture.*
RO	**R**esearch **O**bject *are semantically rich aggregations of resources that bring together data, methods and people in scientific investigations.*
RRegrs	*An eNanoMapper developed tool. It is an R package based on a collection of regression tools from R that can be used to search the best regression models for any dataset. The initial use of the script is aimed at finding QSAR models for chemoinformatics / nanotoxicology.*
SPARQL	**S**imple **P**rotocol **A**nd **R**DF **Q**uery **L**anguage *is a semantic query language for databases, able to retrieve and manipulate data stored in RDF format.*
ToxBank	*A dedicated web-based warehouse for toxicity data management and modelling, a 'gold standards' compound database and repository of selected test compounds, and a reference resource for cells, cell lines and tissues of relevance for in vitro systemic toxicity research carried out across the FP7 HEALTH.2010.4.2.9 Alternative Testing Strategies SEURAT program.*

3.2 ABBREVIATIONS

Abbreviation	Explanation
1D	*1-Dimensional*
3D	*3-Dimensional*
A&A	*Authentication and Authorisation*
AAI	*Authentication and Authorisation Interface*
ANOVA	*Analysis of Variance*
ARFF	*Attribute-Relation File Format*
BAO	*BioAssay Ontology*
bao (namespace)	http://www.bioassayontology.org/bao#
BFO	*Basic Formal Ontology*
BMD	*Benchmark Dose*
BP	*Biological Processes*
CC	*Cellular Components*
CC	*Creative Commons*
CCZero	*Creative Commons Zero waiver*
CDK	*Chemistry Development Kit*
ChEBI	*Chemical Entities of Biological Interest*
CHEMINF	*Chemical Information Ontology*
CL	*Cell Ontology*
clo (namespace)	http://purl.obolibrary.org/obo/
CML	*Chemical Markup Language*

CMS	Content Management System
CSV	Comma-separated values
CV	Cross-Validation
DB	Database
DC	Dublin Core Metadata Ontology
DFT	Density Functional Theory
DoA	Domain of Applicability
DoW	Description of Work
EC	European Commission
EHS	Environment and Health Safety
EJB	Enterprise JavaBeans
EN	Elastic Net
ENET	Elastic Net regression
ENM	Engineered Nano Materials
ENPRA	Risk Assessment Of Engineered Nanoparticles (EU Project)
ENVO	Environment Ontology
EU	European Union
FFNN	Feed Forward Neural Network
GLM	Generalized Linear Model with Stepwise Feature Selection
GNU GPL	GNU General Public License
GO	Gene Ontology
GPW	Gaussian and plane waves
GSEA	Gene Set Enrichment Analysis
GUI	Graphical User Interface
HC	Hierarchical Clustering
HF	Hartree-Fock
HOMO	Highest Occupied Molecular Orbital
HTTP	Hypertext Transfer Protocol
IAO	Information Artifact Ontology
ID3	Iterative Dichotomiser 3
IRI	International Resource Identifier
JPDI	Jaqpot Protocol of Data Interchange JQ Jaqpot Quattro
JRC	Joint Research Centre https://ec.europa.eu/jrc/
JSF	JavaServer Faces
KEGG	Kyoto Encyclopedia of Genes and Genomes
kNN	K Nearest Neighbour
KS-DFT	Kohn-Sham density functional theory
LASSO	Least Absolute Shrinkage and Selection Operator
LC-MS/MS	Liquid chromatography–Mass Spectrometry/ Mass Spectrometry
LDAP	Lightweight Directory Access Protocol
LM	Linear Model
LOO	Leave-One-Out
LOOCV	Leave one out cross validation
LUMO	Lowest Unoccupied Molecular Orbital
MARINA	Managing Risks of Nanomaterials (EU Project)
MeOx	Metal oxides
MF	Molecular Functions

MIREOT	*Minimum Information to Reference an External Ontology Term*
MLR	*Multiple Linear Regression*
MODENA	*Modelling Nanomaterial Toxicity (EU COST Action)*
ModNanoTox	*Modelling nanoparticle toxicity: principles, methods, novel approaches (EU Project)*
MOPAC	*Molecular Orbital PACkage*
MS	*Mass Spectrometry*
MSigDB	*Molecular Signature Databases*
NANOSOLUTIONS	*Biological Foundation for the Safety Classification of Engineered Nanomaterials (ENM): Systems Biology Approaches to Understand Interactions of ENM with Living Organisms and the Environment (EU Project)*
NANoPUZZLES	*Nanoparticles safety collaboration platform*
nanoQSAR	*Nano- Quantitative Structure-Activity Relationship*
NANoREG	*A common European approach to the regulatory testing of Manufactured Nanomaterials (EU Project)*
NCBI	*National Center for Biotechnology Information*
NM	*Nanomaterial*
NN	*Neural Networks*
NPO	*NanoParticle Ontology*
NPs	*Nanoparticles*
nQSAR	*Nano- Quantitative Structure-Activity Relationship*
NSC	*NanoSafety Cluster*
OBI	*Ontology for Biomedical Investigations*
OECD	*Organisation for Economic Co-operation and Development*
OLS	*Ordinary Least Squares*
OpenAM	*OpenAM is an open source access management, entitlements and federation server platform http://en.wikipedia.org/wiki/OpenAM. OpenAM is the software that allows OpenTox services running at each partner facility to use one and the same user credentials (user name and password) this is known as Single Sign On capability.*
OWL	*Web Ontology Language, a family of knowledge representation languages for authoring ontologies*
PATO	*Phenotype and Trait Ontology*
PLS	*Partial Least Squares*
PMML	*Predictive Model Markup Language*
PR	*Protein Ontology (alternative abbreviation)*
PRO	*Protein Ontology*
PubMed	*US National Library of Medicine National Institutes of Health*
QC	*Quality Control*
QM	*Quantum Mechanics*
QSAR	*Quantitative Structure-Activity Relationship*
RA	*Risk Assessment*
RBF	*Radial Basis Function*
RBF-DDA	*Radial Basis Dynamic Decay Adjustment*
RF	*Random Forest*
RFE	*Recursive Feature Elimination*
RMSE	*Root Mean Square Error*
RO	*Relations Ontology*

RO	Research Object
RSS	Rich Site Summary
SAB	Scientific Advisory Board
SDF	Structure Data File
SME	Small and Medium Enterprise
SEO	Search Engine Optimization
SMILES	Simplified molecular-input line-entry system
SMW	Semantic MediaWiki
SOP	Standard Operating Procedure
SPARQL	SPARQL Protocol and RDF Query Language (recursive acronym)
SSL	Secure Sockets Layer
SSO	Single Sign-On
SUN	Sustainable Nanotechnologies (EU Project)
SUNDS	Sustainable Nanotechnologies Decision Support System
SVM	Support Vector Machines
SVN	SubVersioN
SVR	Support Vector Regression
TEM	Transmission Electron Microscopy
TLS	Transport Layer Security
UI	User Interface
UniProt	Universal Protein Database
UO	Unit Ontology
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
UX Design	User eXperience Design
VIP	Very Important Person
VIP	Variable Importance to Projection
WP	Work Package
WS	Web Services
XLS	eXceL Spreadsheet
XML	eXtensible Markup Language
Xvfb	X virtual framebuffer