ENANOMAPPER Report Summary

Project ID: 604134
Funded under: FP7-NMP
Country: Switzerland

Final Report Summary - ENANOMAPPER (eNanoMapper - A Database and Ontology Framework for Nanomaterials Design and Safety Assessment)

Executive Summary:
The eNanoMapper project proposed an ontology and computational infrastructure for toxicological data management of engineered nanomaterials (ENMs) based on open standards, open source, ontologies and an interoperable design to enable a more effective, integrated approach to EU research in nanotechnology. The eNanoMapper concept was built on common components and Application Programming Interfaces (APIs) to address data and models and related interoperability challenges, i.e., providing a formal means for different components to reliably exchange information with each other and to act on it. The flexible computational infrastructure was implemented based on interoperable, standards-compliant and modular web services maximising cross-talk and interaction between different databases. This included key services for ontologies, data storage, data analysis and modelling as well as supporting services (e.g. for authentication and authorisation) and prototype graphical user interfaces (GUIs) for data submission and analysis.

The systems were designed to be interoperable with resources already existing in this area, and where possible eNanoMapper adhered to established standards but also actively contributed to new standards and ontology development.

The developed eNanoMapper infrastructure supports data management in the area of nanosafety research to enable an integrated approach for the risk assessment of nanomaterials. eNanoMapper developed an ontology, a data infrastructure, modelling tools and other resources for predicting toxicity of nanomaterials and worked towards improving the standards in risk assessment of nanomaterials. 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. Further, a collection of descriptors, computational toxicology models and modelling tools were developed, to enable the use and integration of nanosafety data from various sources. The project provided also a rich source of information and documentation (e.g. tutorials, webinars, publications) to support and guide the users.

Project Context and Objectives:
The eNanoMapper project (Grant Agreement 604134; www.enanomapper.net) started in February 2014 with the goal to create a common language in the form of an ontology and computational standards as well as infrastructure for nanomaterials design and safety assessment for the European NanoSafety Cluster (NSC), as well as for the international nanomaterial science community in general. The main challenge of establishing standards and setting up infrastructure lies in the relatively new scientific field of nanomaterial safety: the complexity of toxicology, inherited from chemistry and biology and the additional complexity of characterizing engineered nanomaterials (ENMs) leads to difficulties of ENM identification and hence reproducibility. The result is uncertainty in the validity of experiments and data. This makes it very difficult for scientists, regulators and the industry to share, compare and validate data, models, protocols, and SOPs.

The consortium has assembled partners with specific experience in predictive toxicology, database and modelling resource development (DC, NTUA, IST, IDEA, UM), ontology creation and curation (EMBL-EBI, UM), and existing interactions with
European nanotechnology and NanoSafety Cluster projects (IST, KI, MB and previously VTT). The eNanoMapper project implementation was arranged in seven Work Packages (WPs). The user-driven requirements analysis, design and application testing activities of WP1 provided extremely valuable guidance for all RTD activities. WPs 2, 3 and 4 provided the core component development parts of eNanoMapper in the areas of ontology, data management and computational processing and analysis infrastructure, with WP5 providing a regular integrated testing and release of components including GUls for use application testing and deployment. The use cases guided all components and integrated application development focus and validation by providing software applications to users. WP6 on Dissemination and Training and WP7 on Project Management were pervasive throughout the duration of the project ensuring quality outcomes and results, and relied on the experienced EC project coordination and WP management skills of partners.

Community Outreach (WP1) was responsible for the communication between the eNanoMapper project and the wider nanotechnology community, establishing various mechanisms to ensure a good external collaboration and feedback on the project. The goals were to collaboratively develop interoperability documents together with definitions of personas and use cases that capture the community needs and working practices.

Ontology Development (WP2) aimed to develop and disseminate a comprehensive ontology for the nanosafety domain, encompassing nanomaterials and all information relating to their characterisation, as well as information describing relevant experimental paradigms, biological interactions, safety indications and experimental paradigms.

Database Development & Implementation (WP3) aimed to design, implement and provide the data infrastructure and search capabilities, supporting all aspects of nanomaterials characterization, namely synthesis and processing, physicochemical characterization, environmental and health hazards assessment, high throughput and high content datasets. Implementation of modules and a web service API for different types of queries, namely free text search for protocol description and relevant free text data, SPARQL endpoint for ISA-TAB metadata, and where relevant, a chemical structure and similarity search for pristine and functionalized nanomaterials.

Analysis & Modelling (WP4) ensured that the ontology and data warehouse developed during the project meet the safety-by-design and community needs. One of the main objectives was to develop a computational infrastructure capable to analyse and extract knowledge out of diverse types of nanomaterial-related theoretical descriptors, experimental data and associated metadata, including provenance of experimental data and experimental conditions and protocols.

User Application Development, Integration & Testing (WP5) coordinated the integration, release and testing of eNanoMapper development components into user applications, for the use cases, supporting data uploading, searching, retrieval, analysis, modelling and the creation of reports. These applications were dependent on the infrastructure and services implemented in the other WPs. User and developer feedback was captured in an issue tracker and ticketing system, and used to support development prioritisation and planning.

Dissemination & Training (WP6) ensured an effective dissemination and development of news, findings, progress, lessons learned, practices, resources and services created, and scientific discoveries and inventions from the eNanoMapper project. This was accomplished via the community website (www.enanomapper.net) and through the organization of webinars, workshops, tutorials, as well as the provision of scientific reports and publications.

Management (WP7) ensured an effective planning and management of the project, coordination between partners, administration, and facilitating communications between partners within the project, and between the project consortium and the EC.

eNanoMapper followed an iterative cycle of contextual design and inquiry with users, strategy and specification, which was followed by software development, evaluation and application. The user-centric design, agile software development
methodology and periodic integration, and testing reduced inherent project risk by breaking the WP tasks into smaller sub-
tasks providing more agile ease-of-change across several iterations. Our software source code and build system is visible to the whole NSC community. Across all project development activities, the strategy, concepts, requirements analysis, and design of architecture were developed, tested and refined by an iterative prototyping approach. Various components that resulted from these efforts have been released during the lifetime of the project.

Main objectives

The key goal for eNanoMapper was to enhance scientific validity by creating and establishing standards and by harmonization on different levels:
- To establish various mechanisms for a sound collaboration between eNanoMapper and the nanotechnology/nanosafety community and also to ensure an effective feedback on its objectives;
- To establish a harmonized language with the help of a standardized ontology;
- To accelerate knowledge exchange and reuse through the development of ontologies for the categorisation and characterisation of ENMs in collaboration with other projects;
- To harmonize the scientific processes by establishing an infrastructure for SOPs, protocols or templates, and to harmonize the data sets by setting up a standardized data warehouse, as well as defining standard APIs for the distributed resources of the community;
- To improve the utilisation of data through the implementation of a modular infrastructure for data storage, searching and sharing, based on open standards and semantic web technologies, minimum information standards and established security solutions;
- To develop harmonized and reproducible computational toxicology models, by establishing modelling and API standards;
- To enable the 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;
- To enable the 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;
- To create tools for the exchange, quality assurance and reporting of research protocols and data for regulatory purposes;
- To integrate and test eNanoMapper development components into user applications;
- To disseminate its scientific results, tools and applications to the scientific and technology communities and groups, by providing and facilitating user guidance and interactions;
- To raise awareness of eNanoMapper resources with regulators and policy makers to show their enabling contribution to acceptance of predictive toxicology approaches for nanomaterials safety assessment as alternatives to animal testing;
- To create a community framework to accelerate interdisciplinary collaboration between experimental and computational scientists in the establishment and use of data management and analysis infrastructure and to mutually develop quality-driven guidelines for experimental design and optimal production and sustainable maintenance of new datasets;

Overall, eNanoMapper aimed to support the collaborative safety assessment for ENMs by creating a modular, extensible infrastructure for transparent data sharing, data analysis, and the creation of computational toxicology models for ENMs. Also, based on previous developments of consortium partners in predictive toxicology, biology and nanotechnology research, eNanoMapper aimed to develop resources, tools and standards for a scientifically sound risk assessment of ENMs to support the design of new safe and environment-friendly ENMs as well as the assessment of existing materials. Ultimately, an important goal was to align the eNanoMapper infrastructure development with user needs and to enhance the research cohesion, integration and advancement of the EU NanoSafety Cluster agenda.

Project Results:
During the first phase of the project, the eNanoMapper team carried out a broad requirement analysis, performing detailed interviews with scientists and managers across all main projects of the NSC, including researchers, regulators and industry representatives. Based on the consolidated findings, we were able to prepare use cases, designs and development plans that
were aligned with the cluster’s needs on data, ontology and modelling, and to proceed with work on initial prototypes leading to early releases and integration efforts. The findings were further exploited and used during the whole period of the project to develop, test and implement various solutions towards its primary goal of creating a common language in the form of an ontology, computational standards as well as infrastructure for nanomaterials design and safety assessment.

Community Outreach
The main achievements include reaching out successfully to the EU and US communities, establishment of different collaborations, identification of use cases and establishing an overall systems design. The requirements analysis, which started with interviews, followed by processing into use cases, the derived system design, and the prototypes have ensured that the standards and tools developed will fit with the requirements of the scientists and regulators.

eNanoMapper performed a requirements analysis by interviewing many people involved in the NanoSafety Cluster (NSC) following a contextual design approach. This lead to a systems design and a set of use cases capturing needs from people in the NSC community. Besides dissemination, eNanoMapper has actively been seeking feedback from the community. This is partly done by collaborating with other projects in bilateral activities and by participation in various activities such as harmonization initiatives for modelling and exposure data, international activities around databases, like the NSC WG4 on databases, the US NanoWG, and the US-EU NanoEHS Communities of Research. To capture publicly received feedback, we keep track of online mention of eNanoMapper using CiteULike tool. We have also successfully collaborated with various project on the dissemination of data and knowledge from those other projects. Various solutions are now picked up and considered for adoption. This includes projects such as NANOReg, NanoReg2, caLIBRAtE, OpenRiskNet, ACEnano and others.

Ontology Development
The eNanoMapper ontology covers the full scope of terminology needed to support research in nanomaterial safety. It builds on multiple pre-existing external ontologies such as the NanoParticle Ontology.

We have released an infrastructure for ontology development and maintenance based on a GitHub repository, a Jenkins build environment and also an automated ontology testing software. Further, with an extensive research of the domain area, as well as a survey of existing ontologies, we released the eNanoMapper ontology. We also developed and released as open source library to create slimmed versions of third-party ontologies.

- 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

Prior to building the eNanoMapper ontology, we carried out extensive research of the domain area, as well as a survey of existing ontologies. It was clear that, rather than manually creating each individual term that was required, the eNanoMapper ontology could be compiled in a much more efficient manner by reusing the domain expertise of existing ontologies. To achieve this objective, we developed and released as open source a library to create slimmed versions of third-party ontologies.

Version 1 of the eNanoMapper Ontology was released on 11 March 2015, and contained 7260 terms.
Version 2 of the eNanoMapper Ontology was released on 14 September 2015 and contained 6811 terms.
Most of these have been imported from other ontologies (over 20 separate ontologies have been used), while entirely new terms have been created where no appropriate term could be found in an existing ontology:
- Based on input from various NSC projects, includes classes relating to applications of nanomaterials, drawn from the Chemical Entities of Biological Interest Ontology (CHEBI).
- From the NanoParticle Ontology (NPO), several additional classes of nanoparticles have been added, such as metal nanoparticle and silica nanoparticle.
- Terms from the Statistics Ontology (STATO), the Gene Ontology (GO) and the National Cancer Institute Thesaurus (NCIT)
have been included in the ENM for the first time, providing coverage for several statistical methods, additional assays and study characteristics, respectively.

Version 4 of the eNanoMapper Ontology was made on 26 January 2017. It contains 10,881 terms, an increase of over 2,400 from the previous release:
- 10,295 of the terms in the new release were reused from 22 other ontologies in the biomedical domain.
- The remaining 586 terms were not present in existing ontologies and so were created manually; 321 of these are making their first appearance in the ontology for the final release.

For interactive queries, we have implemented a RDF database based on the Virtuoso triple store and implemented a SPARQL query interface (https://sparql.enanomapper.net). The RDF backend mirrors eNanoMapper ontologies and data and supports combined queries to ontologies and data. For the interactive visualisation of SPARQL query results we have implemented a graphical user interface which is publicly available at https://query.enanomapper.net.

Database Development and Implementation
The main achievements include the eNanoMapper prototype database featuring a flexible data model based on an exhaustive review of existing nano-related entries in chemical- and toxicogenomic-databases. The database REST Application Programming Interface (API) enables user friendly interface and programmatic interaction. The eNanoMapper database is publicly available at http://data.enanomapper.net. It is a public database hosting nanomaterials composition and characterization data, as well as biological and toxicological information. The database content has been continuously updated with datasets provided by project partners and NSC projects. The database software is open source and was recently described in a peer-reviewed publication (doi: 10.3762/bjnano.6.165). Installation instructions are available from http://ambit.sourceforge.net/enanomapper.html. The eNanoMapper database demonstrates the integration of data from multiple sources, using the common data model and API. Evaluations and improvements have been implemented by interactions with the eNanoMapper partners, NSC projects, ISA-TAB, US-Nano WG, publications and public GitHub issue tracker. Ontology annotation was performed with close collaboration with WP2. The import formats currently supported are: OECD Harmonized Templates, NanoWiki RDF and eNanoMapper RDF, and set of spreadsheet templates. A configurable parser was developed to enable the import of the data stored in spreadsheet templates, used by the EU NanoSafety Cluster projects. The export formats have been extended with the new ISA JSONv1 format, following the new ISA specification and with eNanoMapper RDF, based on the BioAssay Ontology data model.

Multiple client implementations for the eNanoMapper database REST API were developed by eNanoMapper partners for different programming languages: Java (by IDEA as part of Ambit codebase ambit.sourceforge.net/source-repository.html and NTUA as part of Jaqpot www.jaqpot.org), Javascript (ambit.js by UM github.com/enanomapper/ambit.js/ and jToxKit by IDEA github.com/ideaconsult/Toxtree.js), R (renm package by UM github.com/enanomapper/renm), Ruby (by IST for Lazar read-across libraries github.com/opentox/lazar); a SPARQL endpoint (sparql.enanomapper.net) and SPARQL query visualizer github.com/enanomapper/enm-ontoviewer). A set of graphical summaries enanomapper.github.io/enmSummaries have been developed using the ambit.js JavaScript client library, demonstrating data access via APIs.

eNanoMapper provides various search functionality: chemically aware search, several options for searching experimental data, semantic search including SPARQL endpoint and free text search, supported by the eNanoMapper ontology and annotated database entries. The free text and faceted search services and the corresponding front end interface at search.data.enanomapper.net were launched Jan 2016 and since then became the most popular user interface to the eNanoMapper database. Importantly, this search interface does a federated search over two different databases (http://data.enanomapper.net and NCI caNanoLab cananolab.nci.nih.gov/caNanoLab/). Feedback from users is gathered through a public issue tracker at GitHub and was taken into account in updated versions of the search application.

Tools for converting the internal model of the eNanoMapper database into ISA-JSONv1 as well as tools for generating ISA-JSONv1 files were developed and released. The export to ISA-JSONv1 is enabled for each data collection and material page in the eNanoMapper database, and most recently enabled in the free text search application. Converting supported input files
into ISA-JSON format is also possible without import into the database and the relevant tools were developed. If needed, the ISA-JSON files can be translated into legacy ISA-TAB. A Substance/Material JSON schema which is a nanomaterial extension of ISA-JSONv1, the counterpart of the ISA-Tab-Nano format. The schema is available at the enanomapper/isa-api fork at GitHub.

D3.3 also reports on the RDF export functionality, allowing data from the AMBIT database to be stored in the SPARQL endpoint. Here the ontology is used to annotate concepts, e.g. to say something is a material, and specific things, e.g. to say something is a specific JRC nanomaterial such as JRCNM01000a. Using this RDF we have implemented two smaller use cases, one about establishing a measure of data quality, and one use case where scientific questions are answered. An example of the latter, summarizing the experimental evidence of a potential genotoxicity of metal oxide nanomaterials was formulated as a SPARQL query. This work was presented as a poster at the international conference SWAT4LS and was presented as an oral presentation at the NMSA conference in Malaga.

We have cleaned, annotated and created configuration files for the publicly released “ISA-Tab-logic“ templates developed by the EC Joint Research Center (JRC) on behalf of NANoREG, enabling direct import into the database through a web browser or programmatically. Being able to read the Excel spreadsheets and write the internal data model into ISA-JSON files accomplishes the goal of automatically generating the ISA files, and enables exporting query results from the database in any desired format. The templates are used or considered for adoption by several NSC projects, such as caLIBRAte and NanoReg2, which are currently implementing the eNanoMapper database data model for collection of data. A documentation site (including templates download) was created accessible via the eNanoMapper tutorials repository at Github [github.com/enanomapper/tutorials/tree/master/DataTemplates](https://github.com/enanomapper/tutorials/tree/master/DataTemplates) and via menu “Data templates” of the search application at [search.data.enanomapper.net](https://search.data.enanomapper.net).

The NANoREG project performed evaluation of the eNanoMapper solutions and in June 2016 took the decision to transfer the data generated by the project to the eNanoMapper database or its successor. Most of the data generated by the NANoREG project is entered via a web entry tool, into a MySQL database, both developed by the Dutch Organization for Applied Scientific Research (TNO). The TNO database design is based on templates developed by the EC JRC for assays performed in NANoREG. The rest of the data is entered into Excel files, following the NANoREG templates, described in the previous sections. A procedure to convert the TNO database and import into the eNanoMapper database was developed and applied on several snapshots of the database, received July, September, October 2016 and January 2017. A test instance of the eNanoMapper database and the search application were installed at [apps.ideaconsult.net/nanoreg1](https://apps.ideaconsult.net/nanoreg1) and [sandbox.ideaconsult.net/search/nanoreg1](https://sandbox.ideaconsult.net/search/nanoreg1). The data access details and a brief user guide were distributed for testing purposes to partners from NANoREG, eNanoMapper and CEINT, Duke University in the US, (upon request from the NANoREG project), and most recently to NanoReg2 partners. On data transfer completion, the NANoREG data will be integrated in the main search application at [search.data.enanomapper.net](https://search.data.enanomapper.net). The status of the NANoREG data transfer to eNanoMapper was described in deliverables, reports, face-to-face meetings and teleconferences with NANoREG. The data upload by NANoREG partners continued until the end of Feb 2017, and the data transfer task will be continued further within the framework of NanoReg2, where the NANoREG data is considered essential for the work planned.

Analysis & Modelling
- Defined the technical specifications of the eNanoMapper analysis and modelling infrastructure. OpenTox algorithm and model APIs were extended to account for the special needs of ENM predictive toxicology and the fact that ENMs are often characterised by a multitude of assays, resulting in high-dimensional datasets.
- Integrated the MOPAC quantum chemistry software for calculating ENM quantum mechanical descriptors from crystallographic data.
- Extended the Java-based Chemistry Development Kit (CDK) with nanomaterial specific descriptors
- Modified and expanded algorithm and modelling OpenTox services in order to exploit the functionalities of the extended API
and able to accommodate the representation of nanoparticle properties in accordance to the eNanoMapper database
- Implemented all major machine learning and data mining algorithms as API compatible web services and used them generate predictive nanoQSAR models for particular use cases, which are exposed to the public as ready to use web applications
- Developed the RRRegs package for computer aided model selection and easy-to-use model comparison. This package automates the estimation of the best regression model by using certain cross validation schemes and by searching over many different algorithms and tuning the parameters in each algorithm
- Developed services for validating the produced predictive models (split, cross and external validation services)
- QSAR Model Reporting Format (QMRF) and QSAR Prediction Reporting Format (QPRF) services were developed and integrated in the eNM modelling infrastructure to summarise information on QSAR/NanoQSAR models and their validation results, based on the REACH objective to reduce animal testing
- Extended work has been done on mechanistic modelling. Particularly, an ENM portal with nanosafety-relevant pathways was initiated to WikiPathways as a basis for pathway analysis, and also workflows for ENMs data analysis, (mainly transcriptomics data) were established, to reveal the significantly and differently affected biological pathways by a variety of exposure scenarios and ENMs (this information was added to the eNanoMapper database)
- Omics data were integrated with biological pathways information to produce a new set of descriptors (named BIO descriptors) for ENMs. An optimisation search functionality estimates the set of descriptors that best predict the toxicity dataset under consideration. Further work on integration of data has been conducted aiming to integrate omics and physicochemical nanoparticles data on a read-across predictive modelling framework. The workflow suggests filtering of the data based on biological pathway information, and then predicts the toxicity index by using either jointly or individually omics and physicochemical ENM data. The produced toxFlow web application is available at http://147.102.86.129:3838/
- Developed an Experimental Design service to serve two main needs of the community, namely a Factorial Design (FD) service for experimental laboratories that would like to design their initial experiments, and an Iterative Experimental Design (IED) service for laboratories with available experimental data aiming to suggest the next experiments or ‘trials’ that need to be conducted
- For the integration of data analysis resources with experimental design, we considered an application to dose-response modelling, after carefully studying the prerequisites with two partners from the SUN NanoSafety Cluster project. A specially designed workflow was the use case based on which an R package was implemented to perform dose-response modelling for continuous data. As an extension, an experimental design functionality was added to the already implemented experimental design services aiming to estimate the optimal number of cases per dosage
- Deployed the Interlaboratory Proficiency Testing service, which comprises a series of calculations and statistical tests in order to assess bias in laboratories and increase repeatability of experiments according to international standards. This is carried out by processing the reported measurements of each laboratory individually against the group consensus
- Developed the jaqpot Quattro (JQ) web application as a platform that integrates ENM modelling and analysis tools. JQ used data from the eNanoMapper database, incorporates the eNanoMapper ontology and is fully compatible with the APIs. The JQ platform has implemented and integrated many of the functionalities mentioned previously: image and quantum mechanical descriptor calculations, data merging and preprocessing, machine learning and data mining algorithms, split, cross and external validation services, interlaboratory proficiency testing techniques, read-across, and optimal experimental design as well as extensive reporting services (including the automatic generation of QPRF reports). The Jaqpot Protocol of Data Interchange (in short JPDI) specifies the form of data exchange between eNanoMapper services and third party algorithm web service implementations and allows developers to integrate their algorithms in the framework. Based on the JDPI protocol, the eNanoMapper computational infrastructure provides wrappers for WEKA, the R language and the Python language. In order to promote the reproducibility of calculations and the ease of dissemination of tools to users, docker images for JQ components have been made available. The capabilities of JQ have been covered thoroughly with a number of documents and video tutorials as well as with the Swagger interface and API documentation hosted at http://jaqpot.org:8080/jaqpot/swagger/. A user interface, enabling non modelling-proficient users to have easy access to all modelling JQ tools and applications, but also with the capability to act as a validated model repository has been developed, available at http://jaqpot.org
- In order to facilitate read across predictions for nanomaterials we extended the lazard read-across framework (https://github.com/opentox/lazar) with capabilities to predict nanoparticle toxicities, additional interfaces for eNanoMapper
data and ontologies and developed a novel method for predicting nanoparticle toxicities based on computed properties alone. These techniques were validated extensively and the results were submitted for publication in the Predictive Toxicology section of Frontiers in Pharmacology. In addition, we have provided the complete source code for lazar libraries, validation experiments and manuscript generation and a docker image with all pre-installed dependencies as an example for reproducible research. With the publication of our complete working environment we intend to encourage other researchers to reproduce and examine our experiments with minimal efforts, and to supply a reference even if system dependencies change in the future.

User Application Development, Integration & Testing
A collaborative working environment was set up for the eNanoMapper project including an issue tracker, a development coordination system and build environment (e.g. Bugzilla issue management system for internal team use and an open source code repository on GitHub including a public issue management). Based on OpenTox services, we set up standard user services as well as libraries for all eNanoMapper components in order to create the base services for all eNanoMapper products for user registration, authentication and authorization. A paper-prototype-specification for key use cases for harmonization and integration purposes was created.

The tools and functionalities development followed the requirements analysis and the prioritized use cases (e.g. user application for importing NanoWiki data, user application for importing NanoSafety Cluster data and user application for free text search of nanomaterials and experimental data in the eNanoMapper and caNanoLab database, user application for conformance to reporting and curation standards, user application on QSAR and read-across for nanomaterials, etc.). Further, 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.

The following user interfaces were developed:
- Jaqpot user interface

The JQ user interface integrates the JQ modelling and analysis tools and is offered to the community as a central modelling platform and application. The JQ can be accessed at [http://www.jaqpot.org/](http://www.jaqpot.org/) and its functionalities include ENM descriptor calculations, nanoQSAR modelling and validation, read-across studies, optimal experimental design and interlaboratory testing (see details under Analysis and Modelling section).

- Nano-lazar user interface

As a frontend for lazar nanoparticle predictions we have developed a web application [https://nano-lazar.in-silico.ch](https://nano-lazar.in-silico.ch) that uses lazar for toxicity predictions and ontologies and data from eNanoMapper for the graphical user interface. The user can enter core and coating structures, physico-chemical properties of nanoparticles or protein interaction data and receives toxicity predictions from a local nano-QSAR (Quantitative Structure-Activity Relationship) model together with a list of similar particles and their physico-chemical and toxicological properties. Links to eNanoMapper ontologies and data as well as to external resources (e.g. BioPortal, UniProt) provide definitions for domain-specific terms and supporting information for toxicity predictions. For this reason, the nano-lazar GUI is also a showcase for the integration of eNanoMapper resources with external services and the functionality of the eNanoMapper API. An additional lazar frontend is provided by nano-lazar REST services, that exposes nano-lazar capabilities via a eNanoMapper compatible API. This interface can be used by external web services and mashups to access nano-lazar functionality.

All nano-lazar libraries and services are tested daily by a process that includes the installation of a complete virtual server from scratch and a comprehensive suite of unit and integration tests. This procedure ensures that all installed components are still compatible and that the communication with external services has not been interrupted (e.g. due to external API changes). 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.

- Knowledge-integrating Applications (Summit)
Web-based applications were created as knowledge-based integrating collaborative support tools for nanosafety related activities such as case study discussions and workshops, with possible extension to other areas. Version 1.0 (https://nanoehs.enanomapper.net/) was used to support the US-EU NanoEHS workshop and its “scrimmage” workshop (Arlington, 2016), while version 2.0 (https://summit.enanomapper.net/) was used to support the EU-US NanoEHS workshop and breakout session (Rheinfelden, 2016). The functionalities of the Summit application facilitated 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 from the workshop.

User Application Documentation

In order to support the developers in this area, eNanoMapper released a broad range of documentation materials and tutorials. Most of the eNanoMapper source code is stored and documented at Github, a web-based repository hosting system with distributed version control and source code management functionality. eNanoMapper project directory on Github at https://github.com/enanomapper, and each repository has an associated issue tracker. For the archival and collecting of different versions of tutorials, there is a separate tutorial repository at https://github.com/enanomapper/tutorials. Also, to facilitate the access to various sources (applications, code documentation, publicly available reports, dissemination materials, etc.) an online publicly available library was set up at http://www.enanomapper.net/library, and we have put together a documentation map available also in GitHub, which offers direct links to all these resources.

Potential Impact:

Impact

eNanoMapper aimed to create a significant positive impact on the nanotechnology industry, regulators and citizens in the European Union and markedly strengthen the cohesive growth, influence and outreach of European nanosafety and research-driven industrial nanotech-based innovation within a global perspective. To achieve these goals, we carried out a broad requirements analysis, performing detailed interviews with scientists and managers across all main projects of the EU NanoSafety Cluster (NSC), including researchers, regulators and industry. Based on the consolidated findings, we were able to prepare use cases, designs and development plans that were aligned with the cluster’s needs on data, ontology and modelling. To build on this, user reviews and acceptances of quality were proposed for ontology, data preparation, input and subsequent use and analysis, in order to successfully show import of data from NSC projects into centralized resources that were compatible with the proposed working standards. On the other hand, for facilitating these collaborations within NSC projects, a simplified procedure should be implemented in the future, in order to support the easier establishment of formal links between projects for data sharing and access (e.g. a Consortium Agreement specification towards supporting cluster level collaboration, data sharing processes, Open license agreements).

To increase the impact on the scientific community, eNanoMapper carried out in the first phase of the project a broad requirement analysis, performing 47 detailed interviews with scientists and managers across all main projects of the EU NSC, including researchers, regulators and industry. Based on the consolidated findings, the use cases were prepared, design and development plans were aligned with the cluster’s needs on data, ontology and modelling, and integration efforts.

Concretely, the eNanoMapper data and ontology platform can now be used to capture data and knowledge along the full lifecycle of ENMs, from research to product development to manufacturing procedures, human and environmental exposures, (eco)toxicological effects and degradation processes. Its flexible design supports the safety, environmental, regulation, and standardisation aspects of ENMs, which are all important for safety-by-design and risk assessment. The linked data approach supports real, achievable and operational interoperability with external ontologies and databases and enables the application of statistical and data mining procedures for data analysis, as well as to support the information transfer activities of the NanoSafety cluster. Thus, these deliverables of eNanoMapper are applicable to exploit and couple together many diverse data types from nanosafety research projects.
In particular, the eNanoMapper platform consists of:

- An ontology addressing all standardisation and regulatory requirements;
- An open source data management software, supporting data sharing and accelerating knowledge flow along the full nanomaterial research and production chain in a secure way (supported by authorization and authentication of all resources);
- Human and machine interfaces for the data warehouse;
- Data analysis platforms supporting toxicological and environmental risk assessment, safety-by-design principles and experimental design;
- Guidelines for experimental design capturing statistical evidence and community-selected standards;
- A set of annotated spreadsheet templates for capturing experimental data and software to convert the spreadsheets into different formats and enabling database import;
- These results have the following impact on the European nanosafety community;
- An agreed language formalized in an ontology which enables comparing and combining private and public data and knowledge;
- An open platform for integrating ENM data sources to provide uniform access to Open and confidential data;
- A platform to support the exploration and highlighting of patterns in ENM structure-activity relations at a US-EU community scale;
- A simplified way of entering data, reducing the cost of safety studies and comparisons;
- The European publishing industry will be provided with clear reporting requirements and usable and well-supported tools, allowing their authors to make their data available to the eNanoMapper platform;
- eNanoMapper approach supports the integration of non-testing methods into risk assessment, and facilitates a harmonised use of existing data and knowledge, enabling finally a significant reduction of animals used for nanomaterials toxicity testing.

Impact on scientific community

eNanoMapper progressed the following community activities:

- Chairing the NSC working groups WG4 databases and WG8 systems biology NSC;
- Leading the NSC Task Force on sustainability;
- Active participation in NSC meetings;
- Contribution to the NSC Database survey;
- Chairing of the US-EU NanoEHS Communities of Research group on Databases & Computational Modelling;
- Contact with the CEN/CENELEC TC 352 nano standards working group;
- Interaction on the ISA-Tab specification with the Oxford ISA-Tab development group;
- Interaction on the ISA-Tab-Nano specification with the US Nano WG;
- Providing advice and comments on the usage and challenges of ISA-Tab-Nano to NSC projects (e.g. NanoPUZZLES);
- Joining the NSC Harmonisation Initiative project group on computational modelling and engaged with this sub-cluster activity on modelling requirements and standards;
- Contributed to a common publication on nanoQSAR led by the NSC Harmonisation Initiative;
- Interactions with the EC-Joint Research Centre on approaches to data templates and ISA-Tab for the NANoREG project;
- Interactions on a data management agreement with the NANOSOLUTIONS project;
- Participated in meetings and information exchange on requirements, templates and data standards with different NSC projects (MARINA, MODENA, SUN, GUIDENano, NanoDefine, NANoREG, NanoPUZZLES, ProSAFE);
- Participated in meetings of and collaborating on various nanoinformatics projects with the US National Cancer Informatics Program Nanotechnology Working Group (Nano WG);
- Participation in the COST action MODENA related activities on modelling of nanomaterials;
- Actively worked with several NSC projects on importing data into the eNanoMapper demonstration data warehouse and on aligning data models (database schema, templates) to the eNanoMapper ontology (NANoREG, ModNanoTox, MARINA, MODENA, NanoPUZZLES);
- Coordination of the discussion on knowledge infrastructure and framework at the level of NanoSafety Cluster and initiating new infrastructure proposals;
Main dissemination and training activities

Detailed descriptions of the dissemination and training activities in years 1, 2, and 3 are provided in the corresponding publicly available deliverables, D6.1, D6.2, and D6.6. Overall, the project has utilized several different strategies to ensure that the project results and resources are communicated to the community. The project partners have been actively engaged in EU Nanosafety Cluster and US-EU nanoEHS activities and members of the consortium chair the working group on databases (WG4) and the working group on systems biology (WG8), and also co-chair the US-EU community of research on databases and computational modeling. The dissemination and training activities were focused on the development of a visual identity and online tools for dissemination and communication including a public webpage for the project, and virtual seminars or webinars as well as online tutorials in order to provide step-by-step guidance on the use of eNanoMapper tools and resources. One dozen tutorials were developed, as reported in the eNanoMapper D6.3 report (available on the eNanoMapper website at www.enanomapper.net). The project has also co-authored a series of publications on different eNanoMapper tools.

eNanoMapper also created web pages in support of other activities including the CompNanoTox conferences in 2015 (Malaga, Spain) and 2017 (Gdansk, Poland) and the 2016 EU-US NanoEHS workshop in Rheinfelden, Germany (2016) as well as an inventory web page for outputs from the NSC modelling projects (www.enanomapper.net/projects/).

We also organized and participated in numerous workshops and conferences in order to disseminate and raise awareness of the eNanoMapper consortium and project applications and to promote discussion and collaborative development and acceptance of shared resources, e.g. ontologies. Project partners are engaged in the elaboration of the roadmap on nanoinformatics, a joint effort between the EU Nanosafety Cluster and US partners. We have discussed with the European Chemicals Agency (ECHA) regarding the adoption of eNanoMapper tools and solutions. Our efforts in promoting community development are summarized in eNanoMapper deliverable report D6.4. During the final year, an exploitation and sustainability plan was prepared (eNanoMapper deliverable report D6.5).

eNanoMapper jointly organized the CompNanoTox Meeting in Benahavis, Spain in 2015 with several other modelling projects in the EU NSC (EU COST action MODENA MembraneNanoPart, ModEnpTox, MODERN, NanoPUZZLES, PreNanoTox) and we have also co-organized the joint nanosafety conference on “New tools and approaches for nanomaterial safety assessment” in Málaga, Spain (2017) together with the FP7 projects, SUN, GUIDENANO, NANOMILE, and NANOSOLUTIONS. We organized the EU-US nanoEHS workshop in Rheinfelden, Germany, in 2016, focused on “Enabling a sustainable harmonised knowledge infrastructure supporting nano environmental and health safety assessment”, and the 2nd Nanosafety Forum for Young Scientists in Visby, Sweden (2016), with the EU Nanosafety Cluster, to foster collaboration between young scientists in the field of nanosafety research.

Exploitation of results

One of the goals of eNanoMapper was to create a collaboration ecosystem in which partners develop specific complementary resources and competencies. Hence, in addition to individual exploitation, the capacity for collaborative exploitation through joint initiatives, tender responses and virtual organisations was explored. This capacity included technical, business, and legal elements and developed further knowledge-oriented collaborative business models. Initially, various options were planned to be explored in order to develop a sustainable financial model, validated against and with the community, including: continued ontology development during future projects, governed by a foundation, business model acting as a hosting provider for proprietary data for stakeholders, business model acting as a hosting provider for calculation services, support contracts helping companies use the ontology, integrating it into custom platforms, selling the developed software along with support contracts (installation, continued development), providing Platform as a Service or Software as a Service models supporting a service-based business ecosystem development. A combination of these options supporting the various components of the outcomes of the project are being pursued.
Another important aim of eNanoMapper was to enhance the impact within the Nano Safety community as envisioned in the description of work. In addition, we elaborated an outline of our plans to extend our outreach, engagement and impact of our key exploitable results. This includes specific considerations of actions with regards to our community engagement, sustainability, training and aligned dissemination activities with the community needs.

The Plan for the Use and the Dissemination of Foreground (PUDF) is structured in two parts: the first describing the strategy and the concrete actions for exploitation of the project’s results and, the second, concerning dissemination. The key exploitable results are included in four cases/categories:

- Data Management systems
- Tools for predictive modelling
- Common language (ontology) for humans and machines
- Integrated risk assessment of nanomaterials

To these results, the following points were addressed: innovativeness, competitive advantages, market size, market trends/public acceptance, product/service positioning, legal or normative or ethical requirements, competitors, prospects/customers, cost of implementation before exploitation, time to market, foreseen product/service price, adequateness of internal staff, external partners to be involved, intellectual property rights, action foreseen, expectation of partners involved, sources of financing. Also the exploitable results priority map was analysed for each of the four cases.

List of Websites:

Web: http://www.enanomapper.net/

Contacts (Name, Affiliation and Email):
Barry Hardy (Douglas Connect GmbH) barry.hardy@douglasconnect.com
Haralampos Sarimveis (National Technical University of Athens) hsarimv@central.ntua.gr
Christoph Helma (In Silico Toxicology GmbH) helma@in-silico.ch
Nina Jeliazkova (IdeaConsult Ltd) jeliazkova.nina@gmail.com
Bengt Fadeel (Karolinska Institutet) bengt.fadeel@ki.se
Gareth Owen (EMBL - European Bioinformatics Institute) gowen@ebi.ac.uk
Egon Willighagen (Maastricht University) egon.willighagen@maastrichtuniversity.nl
Roland Grafstöm (Misvik Biology) grafstromrc@gmail.com