OpenRiskNet

RISK ASSESSMENT E-INFRASTRUCTURE

Workshop

Creating powerful workflows combining data and software services demonstrated on risk assessment case studies

23-24 October 2019 (Amsterdam, The Netherlands)

This project is funded by the European Union

OpenRiskNet: Open e-Infrastructure to Support Data Sharing, Knowledge Integration and *in silico* Analysis and Modelling in Risk Assessment

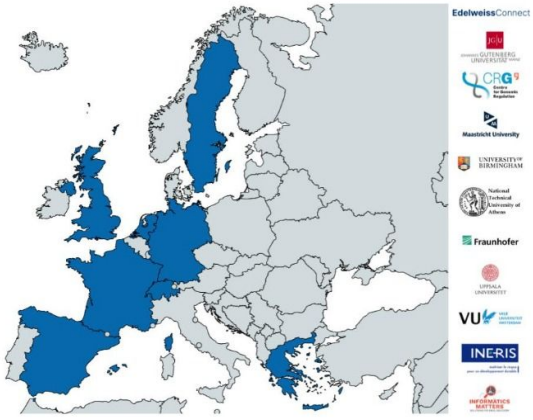
Project Number 731075

www.openrisknet.org

Project identification

Grant Agreement	731075
Project Name	OpenRiskNet: Open e-Infrastructure to Support Data Sharing, Knowledge Integration and *in silico* Analysis and Modelling in Risk Assessment
Project Acronym	OpenRiskNet
Project Coordinator	Edelweiss Connect GmbH
Star date	1 December 2016
End date	30 November 2019
Duration	36 Months

OpenRiskNet Consortium

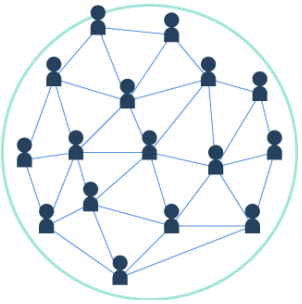


EwC = Edelweiss Connect GmbH, Switzerland
JGU = Johannes Gutenberg-Universität Mainz, Germany
CRG = Fundacio Centre De Regulacio Genomica, Spain
UM = Universiteit Maastricht, Netherlands
UoB = The University Of Birmingham, United Kingdom
NTUA = National Technical University Of Athens, Greece
Fraunhofer = Fraunhofer Gesellschaft Zur Foerderung Der Angewandten Forschung E.V., Germany
UU = Uppsala Universitet, Sweden
IM = Informatics Matters Limited, United Kingdom
INERIS = Institut National De L'environnement Et Des Risques INERIS, France
VU = Vrije Universiteit Amsterdam, Netherlands

Table of Contents

WELCOME NOTE

The **OpenRiskNet** project, developing and implementing an open e-infrastructure addressed to research and industrial communities performing risk assessment of consumer products like chemicals, cosmetics or drugs, is in its third and final year.



To disseminate the achievements and implemented concepts of the project, make the stakeholders aware of the solutions developed, find the best options for the adoption by other projects or organisations, a final workshop is organized. The event will create the perfect environment for current and future users of OpenRiskNet services, developer and administrators as well as the members of related (e-)infrastructure communities to interact with OpenRiskNet developers, modellers and project managers, learn more about the success of the community building efforts fostered by the associated partner programme and the implementation challenges, establish and strengthen links to other projects and organize the transfer of the technology into these.

The workshop is addressing all OpenRiskNet stakeholders (scientific, industrial and regulatory communities) that are invited to participate in this interactive event. This will ensure that all relevant and target groups that need to be aware of the project achievements have access to this information and are enabled to give feedback, and also be trained on the provided solutions.

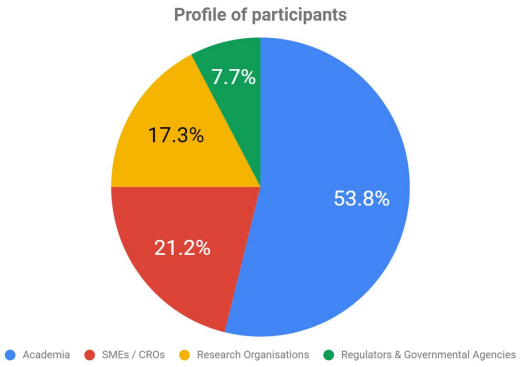
Therefore we invite members of research organisations, academia, regulators and governmental agencies, SMEs and CROs, members of EU infrastructure initiatives to join us for the final workshop. Besides researchers, risk assessors and regulators, the workshop is specifically designed to provide the needed information for developers on how to integrate additional services and for system administrators on deployment options of the virtual research environments in parallel sessions.

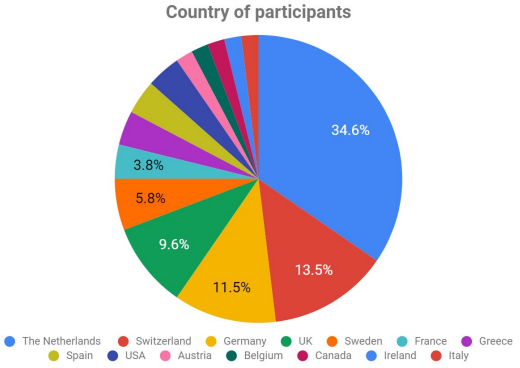
A significant part of the event (day 1) will focus on the **case studies**, demonstrations and hands-on training sessions, where the teams involved in the development and implementation will present and demonstrate the use of the services included in the OpenRiskNet e-infrastructure.

Day 2 will then focus on the outreach and sustainability aspects:

1) The **implementation challenges** were implemented as an innovative part of the **associated partners** programme during the project and linked ten new organisations intensively to project as service providers and early adopters. Complementary services from these partners can now be used and were integrated in the case studies workflow.

2) The **sustainability** measures, the support and maintenance of the infrastructure after the end of the project, collaborations and links created with other communities will be highlighted and discussed with representatives of other major pan-European infrastructures and policy makers.





RESOURCES FOR PARTICIPANTS

- Case studies description: https://openrisknet.org/e-infrastructure/development/case-studies/
- Training materials https://openrisknet.org/library/
- E-infrastructure reference site https://home.prod.openrisknet.org/
- GitHub repository: https://github.com/OpenRiskNet

AGENDA

Day 1 - 23 October 2019

8:30-9:00	*Registration*	
9:00-9:10	**Welcome and introduction to the workshop**	*Thomas Exner (Edelweiss Connect, GmbH, Switzerland)*
9:10-9:30	**OpenRiskNet Case studies introduction and concept**	*Paul Jennings (Vrije Universiteit Amsterdam, The Netherlands)*
9:30-10:30 *8 min / case study*	**OpenRiskNet Case studies** (flash presentations): • **DataCure** - *Thomas Exner (EwC)* • **ModelRX** - *Philip Doganis (NTUA)* • **TGX** - *Danyel Jennen (UM)* • **SysGroup** - *Danyel Jennen (UM)* • **MetaP** - *Daan Geerke (VU)* • **AOPLink** - *Marvin Martens (UM)* • **RevK** - *Philip Doganis (NTUA)*	
10:30-11:00	*Coffee break (incl. posters viewing)*	
OpenRiskNet demo & training sessions		
11:00-12:30	• **Workflows**: practical example of Jupyter notebooks use, Data curation example, workflow across multiple case studies	*Marvin Martens (UM)* *Thomas Exner (EwC)*
12:30-13:30	*Lunch break*	
13:30-15:00	**Parallel sessions**: • **Group 1: Deploying Applications** (*addressed to developers, services providers and infrastructure admins*) (Room: 12W24) • **Group 2: Modelling exercise** (built around the ModelRX case study, support of the DataCure) (*addressed to end-users*) (Room: 01W08)	*Tim Dudgeon (IM)* *Philip Doganis and Pantelis Karatzas (NTUA)* *Denis Gebele (JGU)* *Tomaž Mohorič (EwC)* *Jonathan Alvarsson (UU)*
15:30-16:00	*Coffee break (incl. posters viewing)*	
16:00-17:30	**Parallel sessions**: • **Group 1: Q&A Deploying Applications** (Room: 12W24)	*Tim Dudgeon (IM)*

	• Group 2: Ontology and semantic annotations (Room: 01W08)	*Egon Willighagen (UM)*
18:00-	*Dinner - Buffet organised at workshop venue*	

Day 2 - 24 October 2019

Associated partners session **Chair**: Rex FitzGerald (*Swiss Centre for Applied Human Toxicology, University of Basel, Switzerland*)		
9:00-10:45 *20 min / presentation*	Strategy used to build confidence in **PROSILICO**'s *in silico* methods for prediction of human clinical ADME/PK	*Urban Fagerholm (Prosilico, Sweden)*
	FAME 3: Predicting the Sites of Metabolism in Small Molecules for Phase 1 and Phase 2 Metabolic Enzymes	*Johannes Kirchmair (University of Hamburg, Germany)*
	Using SPARQL to explore human protein data in **neXtProt** and beyond	*Lydie Lane (CALIPHO group, SIB Swiss Institute of Bioinformatics, Switzerland)*
	ToxicoGx: An R platform for integrated toxicogenomics data analysis	*Sisira Kadambat Nair (Princess Margaret Cancer Centre, University Health Network, Toronto, Canada)*
	US EPA AOP-DB: A database resource for the exploration of Adverse Outcome Pathways	*Holly M. Mortensen (US Environmental Protection Agency, Office of Research and Development, USA)*
10:45-11:15	*Coffee break (incl. posters viewing)*	
11:15-12:00	**ToxPlanet:** demo, information on API, discussion on use case scenarios	*Matt Timberlake (ToxPlanet, USA)*
12:00-12:15	**Diamond Light Source:** OpenRiskNet's first external Virtual Environment	*Tim Dudgeon (Informatics Matters Ltd, UK)*
12:15-12:30	**Lessons learned** from the associated partner programme	*Thomas Exner (Edelweiss Connect GmbH, Switzerland)*
12:30-13:30	*Lunch break*	

EU infrastructures & sustainability session **Chair**: Stefan Kramer (Johannes Gutenberg University Mainz, Germany)		
13:30-14:00	**Adoption of OpenRisknet solutions by NanoSafety community and NanoCommons infrastructure**	*Iseult Lynch (University of Birmingham, UK)*
14:00-15:00	1. Introduction to the **European Open Science Cloud (EOSC)** (*10 min*) 2. **OpenAIRE**: an EOSC implementation project (*10 min*)	*Alessia Bardi (Institute of Science and Technology of Information, CNR, Italy)*
	3. **eInfraCentral**: an EOSC implementation project (*10 min*)	*Bjorn Backeberg (EGI Foundation, The Netherlands) on behalf of Jelena Angelis (EFIS Centre, Belgium)*
	4. **EOSC-hub**: an EOSC implementation project (*10 min*)	*Bjorn Backeberg (EGI Foundation, The Netherlands)*
	5. **Technical demo** of cloud and storage services in EOSC (*20 min*)	*Yin Chen (EGI Foundation, The Netherlands)*
15:00-15:30	*Coffee break*	
15:30-16:00	**Panel discussion**: lesson learned and looking ahead to the next steps	
16:00-16:15	Closing session	

ABSTRACTS

Lectures and demonstrations

Adoption of OpenRiskNet solutions by NanoSafety community and NanoCommons infrastructure

Iseult Lynch

School of Geography, Earth and Environmental Sciences, University of Birmingham, UK

Nanomaterials present a unique set of challenges and opportunities for scientists and regulators, arising from quantum confinement (e.g. novel electronic properties, band-gaps that overlap with cells, etc.), and their enormous reactive surface area that leads to agglomeration and binding to biomolecules. Other challenges include the well-known difficulties in terms of batch-to-batch variability in many nanomaterials synthesis or production routes, and the dynamic nature of nanomaterials themselves, whereby ageing can occur during storage and transformations occur upon contact with biological or environmental matrices. Despite these differences, much can be learned from classical toxicology and pharmacology and from the chemoinformatics approaches applied to chemicals. Indeed, it is fully understood that the toxicity pathways and adverse outcomes induced by nanomaterials are not unique to nanomaterials, although the molecular initiating events may be (nano)particle-specific such as surface-generated reactive oxidative species or surface-binding related protein unfolding resulting from corona formation, for example. Thus, the nanosafety community strives to build upon, and leverage, developments from chemoinformatics where possible and adapt for nanomaterials, in the same way it is adapting the regulatory testing guidelines to account for non-equilibrium conditions and non-soluble nanomaterials.

OpenRiskNet aimed to leverage commonalities in approaches across different classes of chemicals, and to develop modular scalable solutions. It focussed on process-based solution, including containerization (e.g. Docker), semantic query language for databases (SPARQL), and integration of a range of chemoinformatics tools for AOP and PBPK modelling via Jupyter notebooks, for example.

Starting as it did 18 months after OpenRiskNet, NanoCommons has been leveraging many of the OpenRiskNet advances, and building on its experiences. For example, OpenRiskNet had to do considerable work around ethics of *in vivo* datasets generated by others and to understand their responsibilities as data hosts / providers. The approaches taken, whereby ethics statements and licence information is included as part of the initial information gathering step from potential dataset providers, and ethics information is included as part of the data and metadata, is adopted in NanoCommons for all datasets curated and imported. OpenRiskNet also developed a searchable Services catalogue approach, allowing the standardised documentation of all services its provides, both those generated by partners

and associated partners as well as services added by external providers, that can be searched and filtered by the category of service, the service type, the user type (developers, end-users, system admin) and the targeted user (e.g. informed public through regulator to data manager). This system, and the services developed, are incorporated into NanoCommons now, as most of the tools developed are also applicable to nanomaterials, or can be further developed as case studies within NanoCommons for use with nanomaterials.

As part of its long-term sustainability planning, NanoCommons is working closely with the European Union Observatory for Nanomaterials (EUON) to understand how best to integrate the NanoCommons Knowledge base and its associated nanoinformatics and risk assessment tools, into the EUON platform, which aims to provide increased transparency regarding the safety of nanomaterials and nano-enabled products. Other sustainability activities underway, co-developed by OpenRiskNet and NanoCommons include working with Elixir to develop a toxicology (including nanotoxicology) community, and more recently to implement Bioschemas annotation for toxicology.

The European Open Science Cloud, its implementation projects (eInfraCentral, OpenAIRE and EOSC-hub) and the EOSC Portal

Alessia Bardi[1] (OpenAIRE), Jelena Angelis[2] (eInfraCentral), Yin Chen[3] (EOSC-hub), Gergely Sipos[3] (EOSC-hub), Björn Backeberg[3] (EOSC-hub)

[1]Institute of Science and Technology of Information, CNR, Italy; [2]EFIS Centre, Belgium; [3]EGI Foundation, The Netherlands

In recent years, the vision of Open Science has emerged as a new paradigm for transparent, data-driven science capable of accelerating competitiveness and innovation. The embodiment of this vision in Europe is the European Open Science Cloud (EOSC), first proposed by the European Commission in April 2016 as part of the Communication on the 'European Cloud Initiative', one of the pillars of the Digital Single Market Strategy. The EOSC puts into practice the European vision for Open Innovation, Open Science and Open to the World by bringing together services and research products such as computing, storage, data, publications, software and workflows from national and international research infrastructures, research performing organisations, collaborations and projects. In so doing, EOSC aims to position Europe as a global leader in scientific data infrastructures and to ensure that European scientists reap the full benefits of data-driven science. The EOSC will offer a virtual environment with open and seamless services for storage, management, analysis and re-use of research data, across borders and scientific disciplines.

Implementation projects such as eInfraCentral, OpenAIRE and EOSC-hub, play central roles in implementing the EOSC vision. The European Commission described them in the Implementation Roadmap for the European Open Science Cloud among the key building blocks of EOSC. The projects bring mature processes, policies and tools from the leading European e-infrastructures to manage the entire life-cycle of services into EOSC, from design to delivery. The entry point into EOSC is the EOSC Portal: http://www.eosc-portal.eu. Here researchers can discover, order, access, use and reuse the services, analytical tools, data

management tools, storage and computing resources they need for their work. Service providers can advertise their services and resources in the portal to make them discoverable and available for access to a broader user group. In this session we present the EOSC vision for open science, the EOSC supporting initiatives, including its implementation projects.

The EOSC Portal is jointly developed and maintained by current EOSC implementation projects: eInfraCentral, EOSC-hub, EOSCpilot and OpenAIRE-Advance.

ToxicoGx - An R platform for integrated toxicogenomics data analysis

Sisira Kadambat Nair[1], Esther Yoo[2], Christopher Eeles[1], Amy Tang[1], Jie Tang[1], Nehme El-Hachem[1], Petr Smirnov[1,2,5], Benjamin haibe-Kains[1,2,3,4,5]

[1]Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada, [2]Department of Medical Biophysics, University of Toronto, [3]Department of Computer Science, University of Toronto, Toronto, M5T 3A1, Canada, [4]Ontario Institute of Cancer Research, Toronto, M5G 0A3, Canada, [5]Vector Institute, Toronto, M5G 1M1, Canada

To better understand the molecular mechanisms underlying compound toxicity, great efforts have been made in screening of drugs/chemicals to generate datasets such as Open-TG GATEs and DrugMatrix. These datasets generate molecular and pharmacological profiles several tissues including liver and kidney along with histopathology assessment and curated literature annotations both in in vivo and in vitro settings. Even though the importance of these datasets is unequivocal, it is important to integrate data across multiple models with maximum overlap and consistency. Some of the major challenges in carrying out comparative analysis of these datasets are the differences in experimental set-up across laboratories and drug annotations, hindering research reproducibility. We have developed ToxicoGx, an open source R package to facilitate integrative analysis of these datasets with intersecting functions that would allow identification of robust biomarkers. A ToxicoSet (TSet) efficiently stores molecular and drug response data along with experimental metadata. It can be easily downloaded and used for data visualization and analyses such as plotting dose response curves, identify differential gene expression induced by drug treatment, and checking gene-compound associations. A detailed tour of the package will be demonstrated in the presentation using knitR (Vignette).

Using SPARQL to explore human protein data in neXtProt and beyond

Lydie Lane and the neXtProt team

SIB Swiss Institute of Bioinformatics and University of Geneva, Geneva, Switzerland

The neXtProt platform (www.nextprot.org) developed at SIB Swiss Institute of Bioinformatics - the Swiss ELIXIR node - proposes solutions to select, explore and reuse available data on human proteins.

The neXtProt team manually curates data from the literature (post-translational modifications, variant phenotypes, protein-protein interactions, etc.) and combines it with

high quality data generated by systems biology projects using a single inter-operable format, the RDF (Resource Description Framework) format. neXtProt data are FAIR (Findable, Accessible, Interoperable, and Reusable), with full traceability ensured by extensive use of metadata.

In the last four years, we have been promoting the use of SPARQL (SPARQL Protocol & RDF Query Language), a semantic query language for databases. Given the wealth of data in neXtProt, our data model is quite complex. neXtProt thus provides over 150 pre-built queries and documentation of its data model to guide the user in his or her first steps. The power and exquisite preciseness of the SPARQL language allow users to simultaneously explore data from neXtProt and data from other semantically compatible resources such as UniProtKB, ChEMBL, DrugBank, Rhea, PDB or WikiPathways.

We would be glad to discuss how we could federate neXtProt data with data from OpenRiskNet resources to benefit risk assessment studies.

Strategy used to build confidence in PROSILICO's in silico methods for prediction of human clinical ADME/PK

Urban Fagerholm and Sven Hellberg

PROSILICO, Sweden

Animal data-based models have been, and still are, the golden standard for predicting ADME/PK and doses of drug candidates in early clinical studies. In vitro models have been developed and are also applied for such predictions. These have apparent advantages and are now accepted as complements to or replacements of animal models. Among limitations of these in vitro systems are a poorer overall predictive accuracy and limited application domain and a belief that simpler systems should not be able outperform those with higher complexity (such as a whole body animal model). Predictive in silico methods have also been developed, and these are mainly applied for screening purposes. Their predictive accuracy and range have not reached those of lab methods. In order to replace lab methods these not only need to show superiority or similar predictive quality, but also a changed mind-set (that a computational model actually can be more reliable than lab methods). PROSILICO has developed an integrated in silico-based platform, including software, for prediction of ADME/PK and doses in humans that outperforms labs in both accuracy and range, and appears to be first to offer a service that can compete with labs. During the talk the strategy used for building this unique predictive platform and confidence in it will be demonstrated.

FAME 3: Predicting the Sites of Metabolism in Small Molecules for Phase 1 and Phase 2 Metabolic Enzymes

Martin Šícho[1,2], Conrad Stork[1], Angelica Mazzolari[3], Christina de Bruyn Kops[1], Alessandro Pedretti[3], Bernard Testa[4], Giulio Vistoli[3], Daniel Svozil[2], Johannes Kirchmair[1,5,6]

[1]Universität Hamburg, Faculty of Mathematics, Informatics and Natural Sciences, Department of Informatics, Center for Bioinformatics, 20146 Hamburg, Germany; [2]University of Chemistry and Technology Prague, Faculty of Chemical Technology, Department of Informatics and Chemistry, CZ-OPENSCREEN: National Infrastructure for Chemical Biology, 166 28 Prague 6, Czech Republic; [3]Università degli Studi di Milano, Facoltà di Scienze del Farmaco, Dipartimento di Scienze Farmaceutiche "Pietro Pratesi", I- 20133 Milan, Italy; [4]Emeritus Professor, University of Lausanne, 1015 Lausanne, Switzerland; [5]University of Bergen, Department of Chemistry, N-5020 Bergen, Norway; [6]University of Bergen, Computational Biology Unit (CBU), N-5020 Bergen, Norway

In silico models for predicting the atoms in a molecule at which metabolic reactions are initiated (sites of metabolism) can help enormously in the development of strategies for the optimization of bioactive compounds. In this contribution we report on the development of the third generation of FAst MEtabolizer (FAME 3), a predictor of sites of metabolism that is based on a collection of extra trees classifiers. FAME 3 is trained on data extracted from the recently launched MetaQSAR database, a comprehensive resource of expert-curated data on xenobiotic metabolism. FAME 3 is one of only a few predictors that cover a wide range of phase 1 and phase 2 metabolic reactions. Benchmark studies on holdout data show that FAME 3 reaches competitive performance. Importantly, FAME 3 features a new, atom-based distance measure that provides an estimate of the reliability of predictions for each atom in a molecule of interest. An API for the OpenRiskNet infrastructure is currently under development.

US EPA AOP-DB: A database resource for the exploration of Adverse Outcome Pathways

Holly M. Mortensen[1], Trevor Levey[2]

[1]US Environmental Protection Agency, Office of Research and Development (ORD), National Health and Environmental Effects Laboratory, Research Triangle Park, NC 27709, USA; [2]Oak Ridge Associated Universities, Research Triangle Park, NC 27709, USA.

There is a need for approaches to understand the biological mechanism of adverse outcomes and human variability in response to environmental chemical exposure. The EPA Adverse Outcome Pathway Database (AOP-DB) is a database resource that combines different data types (AOP, gene, chemical, disease, pathway, orthology, and ontology) to characterize the impacts of chemicals to human health and the environment, and serves as a decision support tool for case study development. Here we present an updated version of the database, which includes an increased number of adverse outcomes and corresponding key events derived from updated feeds from the AOP-Wiki (https://aopwiki.org/), updated disease, phenotype, and ontology information, as well as improved integration with the AOP-DB web user interface. We illustrate these updates and the utility of the AOP-DB through

modules which provide the user with opportunities to explore computationally how chemical-gene molecular key events in adverse outcome pathways link to adverse disease, tissue-specific and population-level health outcomes. This abstract does not reflect EPA Policy.

ToxPlanet

Matt Timberlake

ToxPlanet, USA

Over the last 15+ years, ToxPlanet has assembled and curated hundreds of freely-available high-quality databases available from government agencies, NGO's, and other interested groups. Our AWS-hosted collection, now containing several million documents covering nearly 1 million chemicals, can be accessed using our ORN-compliant API and integrated with additional tools and services. This presentation will include a presentation of the ToxPlanet website, information on our API, and a discussion on current and potential use cases.

Posters

The NanoCommons e-infrastructure – A quick guide to what and how of nanoinformatics in safety assessment suiting basic to expert users in academia, industry, and regulatory agencies

Martin Himly[1], Lucian Farcal[2], Anastasios Papadiamantis[3], Albert Duschl[1], Iseult Lynch[3] and the NanoCommons consortium

[1]University of Salzburg, Austria; [2]Edelweiss Connect GmbH, Basel, Switzerland; [3]University Of Birmingham, United Kingdom;

NanoCommons (https://www.nanocommons.eu/) is a research e-infrastructure project offering access to high quality nanoinformatics services for nanosafety assessors in academia, industry, regulatory agencies. It is user-led, offering and developing the services needed by the user community of nanotechnology, nanosafety and related fields. NanoCommons is built on 3 main pillars: joint research activities, networking activities and transnational access services, covering four categories relevant for nanosafety assessment:

- experimental workflows design and implementation;
- data processing and analysis;
- data visualisation and predictive toxicity;
- data storage and online accessibility.

These services are designed to promote data FAIRness (Findable, Accessible, Interoperable, and Reusable), a key NanoCommons goal, that can be made Open through the NanoCommons Knowledgebase. Thus, NanoCommons provides innovative solutions for data mining, harmonisation, utilisation and re-utilisation, including incorporation of a range of modelling and decision support tools that require and/or can produce organised, high-quality datasets. A number of online training tools have been developed for each of the offered services to help users choose the services relevant to their research questions and applications. The NanoCommons infrastructure encompassing a training library shall bridge academic research with industry and regulators, as recommended by the EU NanoSafety Cluster‘s “Closer to the Market” Research Roadmap serving the Safe-by-Design concept in nanotechnology.

This work is funded by the European Union’s H2020 Research & Innovation Action „NanoCommons – The European Nanotechnology Community Informatics Platform: Bridging data and disciplinary gaps for industry and regulators“ (grant agreement No 731032).

The BridgeDb framework: new functionality and integrations

Egon Willighagen, The BridgeDb Project

Department of Bioinformatics - BiGCaT, Maastricht University, The Netherlands

Biological systems consist of many thousands of genes, proteins, and metabolites. Our collective knowledge about these entities is collected in many hundreds of disparate databases, each focusing on particular themes of knowledge or types of experiments. Importantly, to capture this focus, each database uses its own definition of the key entities and have a matching collection of entity identifiers. Linking all these identifiers requires a flexible tool that provides identifier mappings (IDmaps) but that also knows about these complexities. BridgeDb (http://bridgedb.org) offer IDmaps and understands the difference between a gene, a DNA sequence, a microarray probe, and a protein. Likewise, it understands the difference between stereoisomers, protonation states, and compound classes for metabolites and other small molecules. Modules and plugins are available to use BridgeDb in many bioinformatics toolsets. Furthermore, the latest version fully supports the Linked Open Data Cloud. The BridgeDb project involves developers from Maastricht University, the Netherlands Cancer Institute, and Manchester University, and collaborates with, for example, the Gladstone Institutes, Heriot-Watt University, andUppsala University.

A network-based approach for predicting key events and transcription regulators associated with nanomaterial-induced toxicity

Vadim Zhernovkov[1], Tapesh Santra[1], Hilary Cassidy[1], Oleksii Rukhlenko[1], David Matallanas[2], Aleksandar Krstic[1], Walter Kolch[1,2,3], Vladimir Lobaskin[4], and Boris N. Kholodenko[1,2,3,5]

[1]Systems Biology Ireland, University College Dublin, Dublin 4, Ireland; [2]School of Medicine and Medical Science, University College Dublin, Belfield, Dublin 4, Ireland; [3]Conway Institute of Biomolecular & Biomedical Research, University College Dublin, Ireland; [4]School of Physics, University College Dublin, Dublin 4, Ireland; [5]Department of Pharmacology, Yale University School of Medicine, New Haven CT 06520, USA

A rapid increase of new nanomaterial (NM) products poses new challenges for their risk assessment. Current traditional methods for estimating potential adverse health effect of NMs are complex, time consuming, and expensive. In order to develop new prediction tests for nanotoxicity evaluation, a systems biology approach, and data from high-throughput omics experiments can be used. We present a computational approach that combines reverse engineering techniques, network analysis and pathway enrichment analysis for inferring the transcriptional regulation landscape and its functional interpretation. To illustrate this approach, we used published transcriptomic data derived from mice lung tissue exposed to carbon nanotubes (NM-401 and NRCWE-26). Because fibrosis is the most common adverse effect of these NMs, we included in our analysis the data for bleomycin (BLM) treatment, which is a well-known fibrosis inducer. We inferred gene regulatory networks for each NM and BLM to capture functional hierarchical regulatory structures between genes and their regulators. Despite the different nature of the lung injury caused by nanoparticles and BLM, we identified several conserved core regulators for all agents. We reason that these regulators can be considered as early predictors of toxic responses after

NMs exposure. This integrative approach, which refines traditional methods of transcriptomic analysis, can be useful for prioritization of potential core regulators and generation of new hypothesis about mechanisms of nanoparticles toxicity.

Development of an API for the Site of Metabolism Predictor FAME 3

Martin Šícho[1,2], Daniel Svozil[2], Johannes Kirchmair[1,3,4]

[1]Faculty of Mathematics, Informatics and Natural Sciences, Department of Informatics, Center for Bioinformatics, UniversitätHamburg, 20146 Hamburg, Germany; [2]Faculty of Chemical Technology, Department of Informatics and Chemistry, CZ-OPENSCREEN: National Infrastructure forChemical Biology, University of Chemistry and Technology Prague, 166 28 Prague 6, Czech Republic; [3]Department of Chemistry and Computational Biology Unit (CBU), University of Bergen, N-5020 Bergen, Norway; [4]University of Bergen, Computational Biology Unit (CBU), N-5020 Bergen, Norway

An important factor to consider in risk assessment is the metabolic fate of xenobiotics. In this endeavor, the availability of a software tool to accurately predict sites of metabolism (atoms in a molecule at which metabolic reactions are initiated) would help to speed up the process of understanding a compound's metabolism and dramatically reduce experimental costs. FAME 3 [1] is the last iteration of FAst MEtabolizer, a site of metabolism predictor utilizing random forest classifiers and atom descriptors. Based on a large expert-curated data set (MetaQSAR [2]), FAME 3 is one of few predictors that cover both phase 1 and phase 2 metabolic reactions and, unlike other predictors, features an estimate of the reliability of individual SoM predictions. In this work, we briefly describe the relevant features of FAME 3, but more importantly we focus on the specification and usage examples of the FAME 3 application programming interface (API) that is currently being implemented within the OpenRiskNet infrastructure. FAME 3 and its OpenRiskNet API are free for academic and noncommercial research, and this work should serve as a guide to anyone who would like to use this software for not only risk assessment, but also in any other context.

References

1. FAME 3: Predicting the Sites of Metabolism in Synthetic Compounds and Natural Products for Phase 1 and Phase 2 Metabolic Enzymes Martin Šícho, Conrad Stork, Angelica Mazzolari, Christina de Bruyn Kops, Alessandro Pedretti, Bernard Testa, Giulio Vistoli, Daniel Svozil, and Johannes Kirchmair Journal of Chemical Information and Modeling 2019 59 (8), 3400-3412 DOI: 10.1021/acs.jcim.9b00376
2. MetaQSAR: An Integrated Database Engine to Manage and Analyze Metabolic Data Alessandro Pedretti, Angelica Mazzolari, Giulio Vistoli, and Bernard Testa Journal of Medicinal Chemistry 2018 61 (3), 1019-1030 DOI: 10.1021/acs.jmedchem.7b01473

OpenRiskNet Part I: Development of an open e-infrastructure predictive toxicology and risk assessment

Thomas Exner[1], Lucian Farcal[1], Daniel Bachler[1], Nofisat Oki[1], Denis Gebele[2], Atif Raza[2], Stefan Kramer[2], Evan Floden[3], Cedric Notredam[3], Jordi Rambla[3], Danyel Jennen[4], Jumamurat Bayjanov[4], Egon Willighagen[4], Marvin Martens[4], Chris Evelo[4], Iseult Lynch[5], George Gkoutos[5], Philip Doganis[6], Pantelis Karatzas[6], Haralambos Sarimveis[6], Marc Jacobs[7], Ola Spjuth[8], Tim Dudgeon[9], Alan Christie[9], Frederic Bois[10], Daan Geerke[11], Paul Jennings[11], Barry Hardy[1]

[1]Edelweiss Connect GmbH, Basel, Switzerland; [2]Johannes Gutenberg-Universität Mainz, Germany; [3]Fundacio Centre De Regulacio Genomica, Spain; [4]Maastricht University, Netherlands; [5]University Of Birmingham, United Kingdom; [6]National Technical University Of Athens, Greece; [7]Fraunhofer Gesellschaft, Germany; [8]Uppsala Universitet, Sweden; [9]Informatics Matters Ltd., United Kingdom; [10]Institut National De L'environnement Et Des Risques, France; [11]Vrije Universiteit Amsterdam, Netherlands.

OpenRiskNet (https://openrisknet.org/) is a 3-year project funded by the EU within Horizon 2020 EINFRA-22-2016 Programme, with the main objective to develop an open e-infrastructure providing data and software resources and services to a variety of industries requiring risk assessment (e.g. chemicals, cosmetic ingredients, pharma or nanotechnologies).

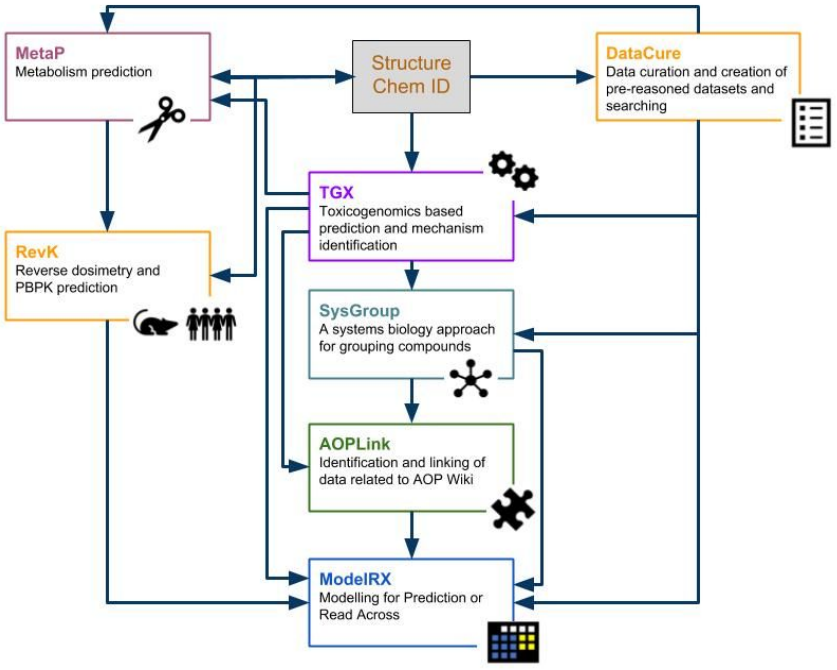


Figure 1. Structure of interlinked case studies targeting different parts of the risk assessment workflow

The infrastructure is built on virtual research environments (VREs), which can be deployed to workstations as well as public and in-house cloud infrastructures. Services providing data, data analysis, modelling and simulation tools for risk assessment are integrated into the

e-infrastructure and can be combined into workflows using harmonised and interoperable application programming interfaces (APIs) (https://openrisknet.org/e-infrastructure/services/). For complete risk assessment and safe-by-design studies, OpenRiskNet e-infrastructure functionality is combined via a variety of incorporated services demonstrated within a set of case studies (see figure 1). The case studies present real-world settings such as data curation, systems biology approaches for grouping compounds, read-across applications using chemical and biological similarity, and identification of areas of concern based only on alternative methods (non-animal testing) approaches.

OpenRiskNet is working with a network of partners, organised within an Associated Partners Programme, aiming to strengthen the working ties to other organisations developing relevant solutions or tools.

OpenRiskNet Part II: Predictive Toxicology based on Adverse Outcome Pathways and Biological Pathway Analysis

Marvin Martens[1], Thomas Exner[2], Nofisat Oki[2], Danyel Jennen[1], Jumamurat Bayjanov[1], Chris Evelo[1], Tim Dudgeon[3], Egon Willighagen[1]

[1]Maastricht University, Netherlands; [2]Edelweiss Connect GmbH, Basel, Switzerland; [3]Informatics Matters Ltd., United Kingdom.

The OpenRiskNet project (https://openrisknet.org/) is funded by the H2020-EINFRA-22-2016 Programme.

Here we present how the concept of Adverse Outcome Pathways (AOPs), which captures mechanistic knowledge from a chemical exposure causing a Molecular Initiating Event (MIE), through Key Events (KEs) towards an Adverse Outcome (AO), can be extended with additional knowledge by using tools and data available through the OpenRiskNet e-Infrastructure. This poster describes how the case study of AOPLink, together with DataCure, TGX, and SysGroup, can utilize the AOP framework for knowledge and data integration to support risk assessments. AOPLink involves the integration of knowledge captured in AOPs with additional data sources and experimental data from DataCure. TGX feeds this integration with prediction models of the MIE of such AOPs using either gene expression data or knowledge about stress response pathways. This is complemented by SysGroup, which is about the grouping of chemical compounds based on structural similarity and mode of action based on omics data. Therefore, the combination of these case studies extends the AOP knowledge and allows biological pathway analysis in the context of AOPs, by combining experimental data and the molecular knowledge that is captured in KEs of AOPs.

OpenRiskNet Part III: Modelling Services in Chemical/Nano-safety, Environmental Science and Pharmacokinetics

Stefan Kramer[1], Philip Doganis[2], Denis Gebele[1], Atif Raza[1], Pantelis Karatzas[2], Haralambos Sarimveis[2], Jonathan Alvarsson[3], Ola Spjuth[3], Staffan Arvidsson[3], Thomas Exner[4], Lucian Farcal[4], Barry Hardy[4] for the OpenRisknet Consortium

[1]Johannes Gutenberg-Universität Mainz, Germany; [2]National Technical University Of Athens, Greece; [3]Uppsala Universitet, Sweden; [4]Edelweiss Connect GmbH, Basel, Switzerland

The OpenRiskNet project (https://openrisknet.org/) is funded by the H2020-EINFRA-22-2016 Programme and its main objective is the development of an open e-infrastructure providing data and software resources and services to a variety of industries requiring risk assessment (e.g. chemicals, cosmetic ingredients, pharma or nanotechnologies).

The concept of case studies was followed in order to test and evaluate proposed solutions and is described in https://openrisknet.org/e-infrastructure/development/case-studies/. Two case studies, namely ModelRX and RevK, focus on modelling within risk assessment.

The ModelRX – Modelling for Prediction or Read Across case study provides computational methods for predictive modelling and support of existing data suitability assessment. It supports final risk assessment by providing calculations of theoretical descriptors, gap filling of incomplete datasets. computational modelling (QSAR) and predictions of adverse effects. Services are offered through Jaqpot (UI/API), JGU WEKA (API), Lazar (UI) and Jupyter & Squonk Notebooks.

In the RevK – Reverse dosimetry and PBPK prediction case study, physiologically based pharmacokinetic (PBPK) models are made accessible for the purpose of risk assessment-relevant scenarios. The PKSim software, the httk R package and custom-made PBPK models have been integrated. RevK offers services through Jaqpot (UI/API).

OpenRiskNet Part IV: WEKA Machine Learning Services for the Prediction of Half-Lifes of Chemicals and Nanoparticle Transport

Stefan Kramer, Denis Gebele, Atif Raza

Johannes Gutenberg-Universität Mainz, Germany

The OpenRiskNet project (https://openrisknet.org/) is funded by the H2020-EINFRA-22-2016 Programme and its main objective is the development of an open e-infrastructure providing data and software resources and services to a variety of industries requiring risk assessment (e.g. chemicals, cosmetic ingredients, pharma or nanotechnologies).

We will present the WEKA machine learning services within the infrastructure and how they can be used to solve complex prediction tasks: the prediction of (i) half-life of chemicals under given environmental conditions and of (ii) nanoparticle transport behavior from physicochemical properties. For that purpose, we will reconstruct previous efforts using complex workflows and architectures and simplify the models while maintaining their prediction performance. In both cases, the overall problem (predicting the fate of a

compound depending on its properties and external conditions) is modeled as a cascaded prediction model, where the prediction of one model is, with particular attention to validity and performance, entering another model as input. The approach performs well on the half-life data, while the nanoparticle data are too noisy and incomplete to warrant more than the most basic models. Overall, the reconstruction of the two applications within OpenRiskNet provides more evidence for the power and versatility of the framework.

LOGISTICS

Website and registration: https://openrisknet.org/events/74/
Contact: openrisknet@edelweissconnect.com

Venue

Vrije Universiteit, Auditorium, O|2 Human Life Sciences building, De Boelelaan 1108, 1081 HZ Amsterdam, The Netherlands

https://goo.gl/maps/eoeYMjUStZqGEX437

Recommended hotels

- Novotel Amsterdam City https://www.accorhotels.com/gb/hotel-0515-novotel-amsterdam-city/index.shtml
- Motel One Amsterdam https://www.motel-one.com/en/hotels/amsterdam/hotel-amsterdam/
- Amsterdam Forest Hotel https://www.amsterdamforesthotel.com/
- Spa Sport Hotel Zuiver https://zuiveramsterdam.nl/hotel
- Olympic Hotel http://olympic-hotel.hoteleamsterdam.net/en/
- Ammonite Hotel Amsterdam https://www.ammonite-hotel.com/en/index.html