OpenRiskNet, an Open e-Infrastructure to Support Data Sharing, Knowledge Integration and In Silico Analysis and Modelling in Risk Assessment

Thomas Exner (DC), Joã Todorík (DC), Daniel Bachler (DC), Lucian Faraci (DC), Chris Ewel (UM), Egon Willighagen (UM), Marc Jacobs (Fraunhofer), Filip Degrieck (NTUA), Haralampos Sambivsw (NTUA), Ioannis Zampounoglou (UoB), Stefan Kraemer (GGU), Cetin Nektarbay (CRG), Daniel Jersey (UM), Georgios Glotsios (UM), Paul Jennings (UO), Tim Douglan (UM), Frederick Bins (INESR) and Barry Hardy (UM)

Introduction

OpenRiskNet is a 3-year project funded by the European Union’s Horizon 2020 program. OpenRiskNet is developing an open e-infrastructure to provide access to resources and services to a variety of communities requiring risk assessment (chemicals, cosmetic ingredients, therapeutic agents, nanomaterials, etc.).

Toxicology and risk assessment are undergoing a paradigm shift, from a phenomenological to a mechanistic discipline based on in vitro and in silico approaches. These new approaches represent an important alternative to classical animal testing applied to the evaluation of chronic and systemic toxicity risks. Large databases and highly sophisticated methods, algorithms and tools are available for tasks such as hazard prediction, toxicokinetics and in vitro in-situ extrapolations to support this transition. However, since these services are not interoperable and independently provided by different groups, there is no standardized way to access the data or run modelling workflows. To overcome the fragmenation of services and tools, OpenRiskNet will provide concepts for harmonized and interoperable open e-infrastructure resources and services.

OpenRiskNet combines the achievements from earlier projects in generating and validating workflows, knowledge integration and data access.

The main components of the infrastructure will be an interoperability layer based on well-defined application programming interfaces (APIs) added to every service to describe the functionality and guaranteed technical and semantic interoperability, a discovery service, deployment options based on container technology, and packaging of the infrastructure into virtual instances.

API Design Concept

It is necessary to go beyond API concepts realised in previous projects in this area (OpenTox and Open PHACTS) for two reasons:
1. A much broader scope of data and tool will be integrated and harmonised in OpenRiskNet; and
2. Semantic annotation of APIs needs to be provided, with more information on the service, input and output.

OpenRiskNet will make the interfaces smarter by adding a semantic interoperability layer. By providing this layer, a service will provide the following information to be compliant with the OpenRiskNet infrastructure:

- Scientific background of the service
- For example, a link to the relevant publication but also to manuals, tutorials and other training materials
- Technical background
- Links to source code, installation instructions, license information and deployment options
- Capabilities of the service which can be chosen by the user to optimise the results
- For databases, this will include, among other things, the used schema, i.e. the description of the stored data and the associated metadata, as well as search options; for software tools, it includes the type and amount of generated output including the options and parameters
- Requirement on input data types and formats and options on the output format.

The approach to creating the semantic layer that looks promising is a combination of OpenAPI definitions with SDLC data workflows to bridge the worlds of API development and the semantic web. Below is a simple example of how such a description produces a hierarchical presentation of a dataset.

Case Studies

Case Studies are used to:
- Test and evaluate the solutions provided, especially regarding the usability of the developed APIs and the interoperability layer.
- Demonstrate the ability to satisfy the requirements of diverse stakeholder groups (researchers, risk assessors and regulators).
- Present real-world applications like:
  - Systems biology approaches for grouping compounds and
  - Read-across applications using chemical and biological similarity.
- Identify areas of concern based on in vitro and in silico approaches for compounds lacking any previous knowledge from animal experiments (ab initio case).
- Go to the prioritization of data sources and tools to be integrated and used as first examples to improve the level of the corresponding APIs with respect to harmonization of the API endpoints, service description and semantic annotation.

A workflow for the safety assessment of chemicals without animal testing developed within the SEURAT-1 initiative (Bergen et al., 2017) was selected to guide the initial definition of the case studies.

This workflow constructs a hypothesis based on existing data, computer models, biomarker considerations, and then, targeted non-animal testing.

For further developments, other concepts will be add or followed in order to avoid the limitations of a single framework, and this information will be incorporated into the related case study description.

Data API for ToxCast and TG-Gates

Harmonisation of APIs for data and computational services opens many opportunities to combine the tools in new and innovative ways, build workflows to answer complex questions and to develop more efficient ways to find and use data across multiple sources. If the existing first version of harmonized data APIs for ToxCast and TG-Gates is not semantically annotated, it already allows the search for compounds and endpoints over multiple datasets using the data explorer and the easy integration of the data in workflow tools like KNIME.

References and More Information


Benefits for associated partners

Service providers: greater visibility of their tools by being listed in the OpenRiskNet discovery service. "Infinite" additional features by combining with other tools, support for emerging techniques like API development and implementation/depolyment

Early adopters: easy access to a increasing number of tools using their preferred access route (web, workflow tools (like KNIME, scripts)) without the need of manually downloading of data and file conversion when moving from one tool to another, harmonized access for comparison of different approaches.

Technology partners: getting feedback of the usability of the services and tools, seeing them in a real-world application and being able to use the success story for marketing.

If you are interested, please contact Thomas Exner at thomas.exner@dsouglassconnect.com or visit https://openrisknet.org/associated-partner-programme/