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

OpenRiskNet

douglasconnect working communities

RISK ASSESSMENT E-INFRASTRUCTURE

Thomas Exner (DC), Joh Dokler (DC), Daniel Bachler (DC), Lucian Farcal (DC), Chris Evelo (UM), Egon Willighagen (UM), Marc Jacobs (Fraunhofer), Philip Doganis (NTUA), Haralambos Sarimveis (NTUA), Iseult Lynch (UoB), Stefan Kramer (JGU), Cedric Notredame (CRG), Danyel Jennen (UM), Georgios Gkoutos (UoB), Ola Spjuth (UU), Paul Jennings (VU), Tim Dudgeon (IM), Frederic Bois (INERIS) and Barry Hardy (DC)

Thomas Exner I thomas.exner@douglasconnect.com I +49 171 3807485

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 resources and services to a variety of communities requiring risk assessment (chemicals, cosmetic ingredients, therapeutic agents, nanomaterials, et al.).

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 vivo* extrapolations to support this transition. However, since these services are developed independently and provided by different groups worldwide, there is **no standardized way to access the data or run modelling workflows**. To overcome the fragmentation of data 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 modelling and validation workflows, knowledge integration and data.

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 guaranteeing **technical and semantic interoperability**, a discovery service, deployment options based on container technology, and packaging of the infrastructure into virtual instances.
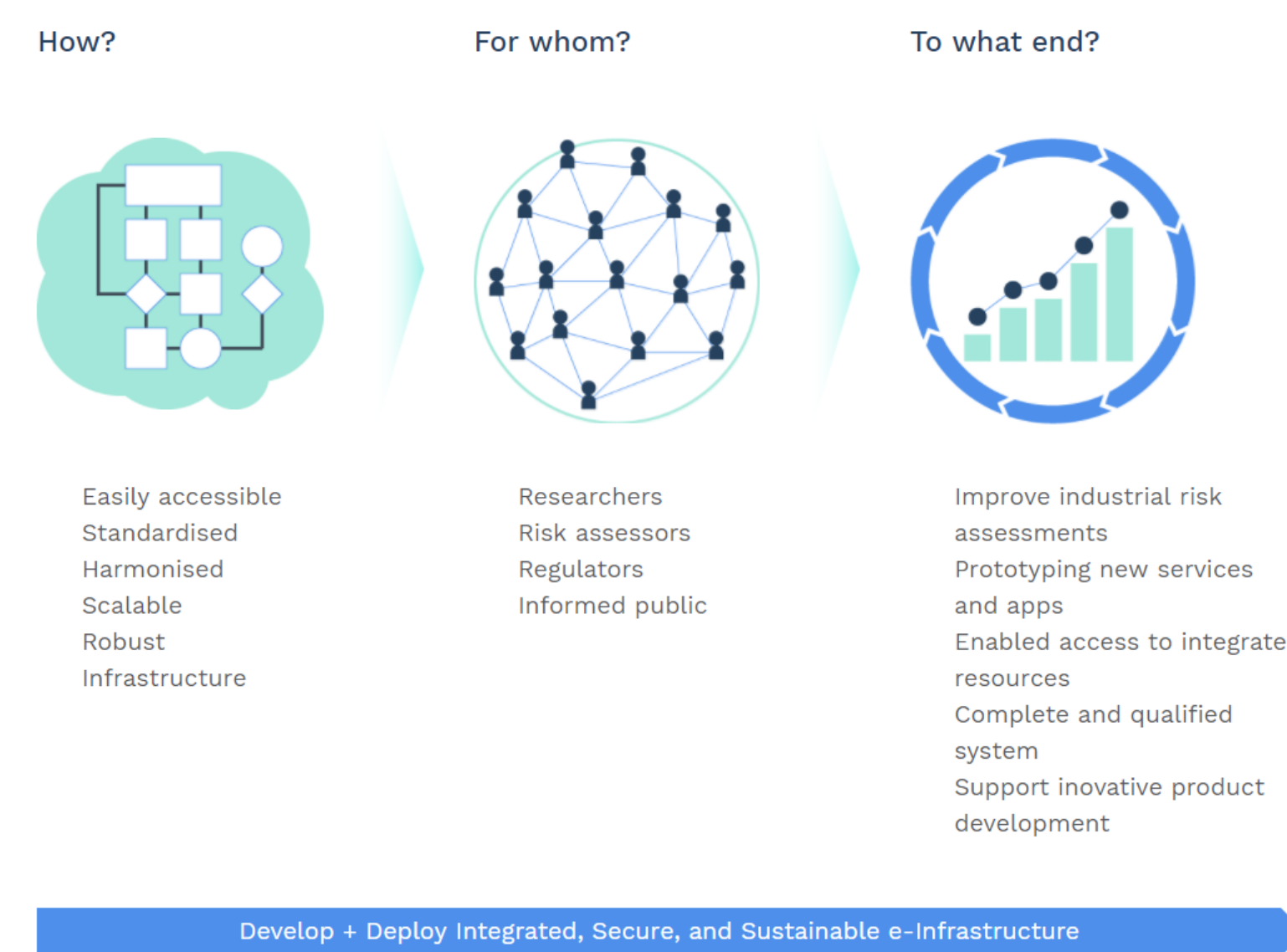
API Design Concept

It is necessary to go beyond API concepts realized 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 harmonized 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 querying 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 optimize the results
  - For databases, this will include, among other things, the used data schemata, 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 JSON-LC data serialization 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.

JSON-LD Input
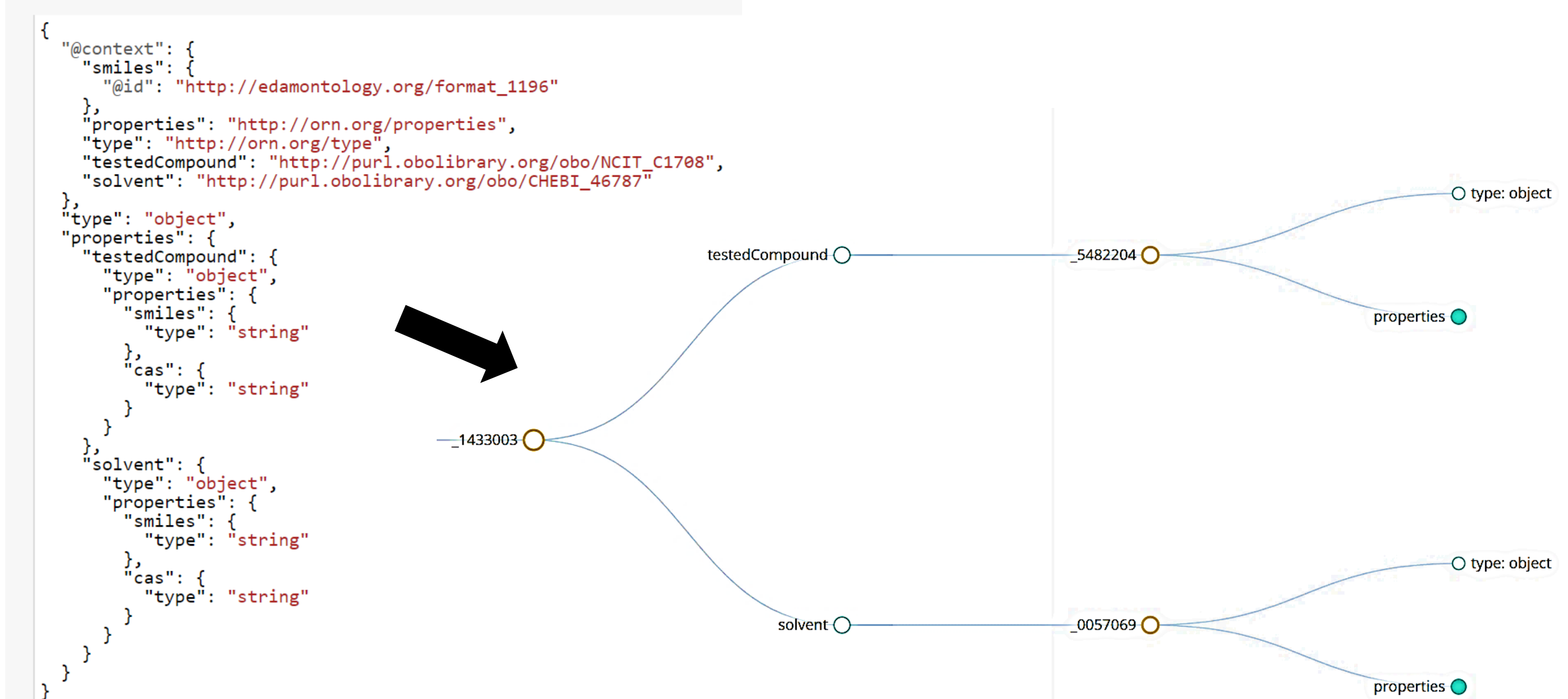
Case Studies

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).
- Guide 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** (Berggren et al., 2017) was selected to guide the initial definition of the case studies.

This workflow constructs a **hypothesis based on existing data, computational modelling, biokinetic considerations, and then, targeted non-animal testing**.

For further developments, other concepts will be added or followed in order to avoid the limitations of a single framework, and this information will be incorporated into the related case study description.
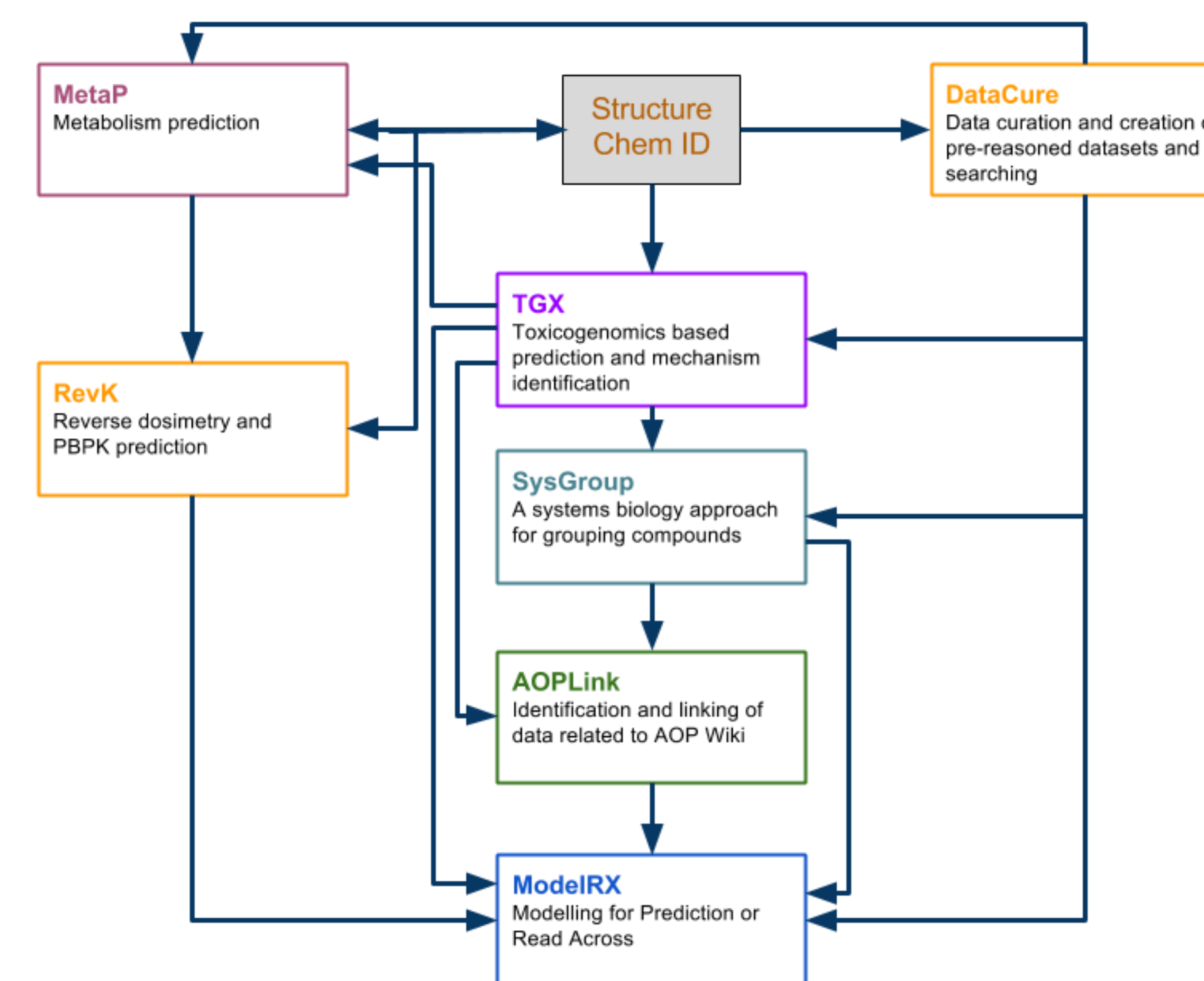
Data API for ToxCast and TG-Gates

Harmonization 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**. Even 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.
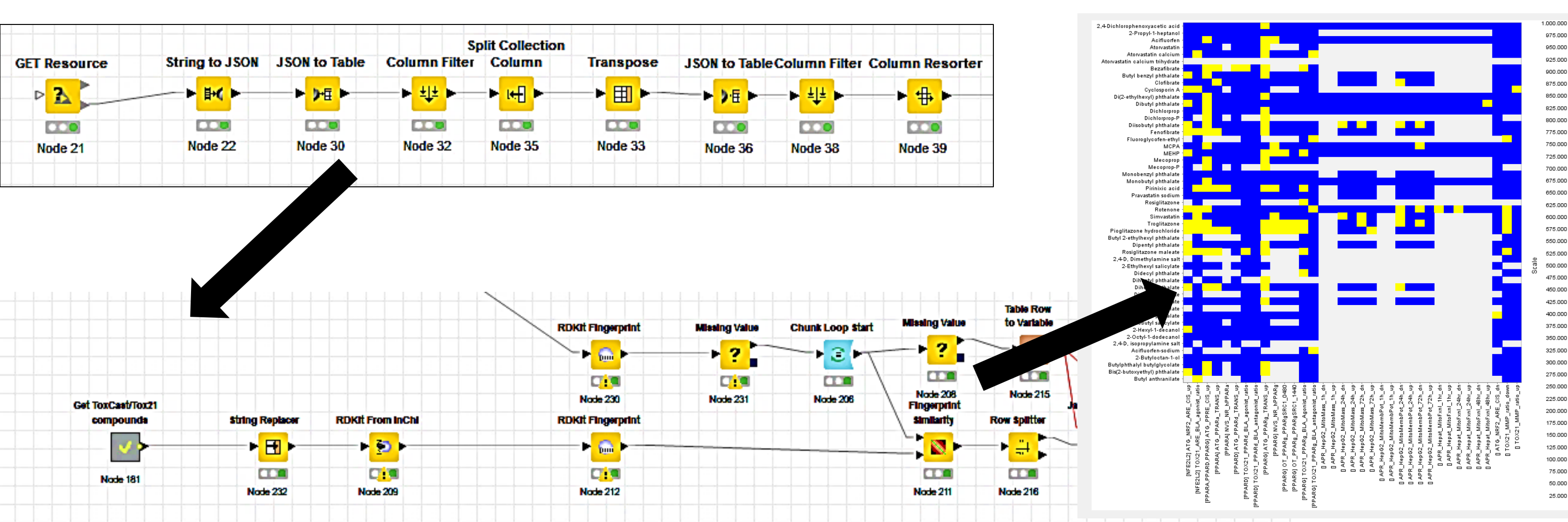

OpenRiskNet Partners

P1   Douglas Connect GmbH, Switzerland (DC)
P2   Johannes Gutenberg-Universitat Mainz, Germany (JGU)
P3   Fundacio Centre De Regulacio Genomica, Spain (CRG)
P4   Universiteit Maastricht, Netherlands (UM)
P5   The University Of Birmingham, United Kingdom (UoB)
P6   National Technical University Of Athens, Greece (NTUA)
P7   Fraunhofer Gesellschaft Zur Foerderung Der Angewandten Forschung E.V., Germany (Fraunhofer)
P8   Uppsala Universitet, Sweden (UU)
P9   Medizinische Universitat Innsbruck, Austria (MUI)
P10 Informatics Matters Limited, United Kingdom (IM)
P11 Institut National De L'environnement Et Des Risques INERIS, France (INERIS)
P12 Vrije Universiteit Amsterdam, Netherlands (VU)

Associated Partner Program

To ensure the **usability of the infrastructure, alignment with the community as well as to pursue complete coverage of important tools**, OpenRiskNet works with a network of partners, organized in the Associated Partners Program. It was officially launched at the OpenTox Euro 2017 conference.

The Program aims at strengthening the working ties between the OpenRiskNet Consortium members and other organizations within relevant scientific and technology communities. Any organization such as a university, institute, consortium, non-governmental organizations (NGOs), as well as small and medium enterprises (SMEs) or large commercial companies can become an Associate Partner of OpenRiskNet.

We expect to have different types of users:
1) **Service providers**, who want to integrate their databases and software tools into the OpenRiskNet infrastructure,
2) **Early adopters**, who will use the infrastructure for their predictive toxicology and risk assessment tasks, and
3) **Technology partners**, who develop services and tools on which the OpenRiskNet e-infrastructure will be based (e.g. OpenShift by Red Hat).

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 containerization/deployment

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@douglasconnect.com or visit https://openrisknet.org/associated-partner-programme/.

References and More Information

Elisabet Berggren, Andrew White, Gladys Ouedraogo, Alicia Paini, Andrea-Nicole Richarz, Frederic Y. Bois, Thomas Exner, Sofia Leite, Leo A. van Grunsven, Andrew Worth, Catherine Mahony, "Ab initio chemical safety assessment: A workflow based on exposure considerations and non-animal methods", Computational Toxicology, Volume 4, 2017, Pages 31-44.

https://openrisknet.org/development/api-concept/
https://openrisknet.org/development/case-studies/
https://json-ld.org/spec/latest/json-ld/
https://www.openshift.com/

Acknowledgements