AI-assisted Next Generation Risk Assessment and Safe and Sustainable Design Workflows enabled by FAIR Data and Knowledge

Barry Hardy, Daniel Burgwinkel, Tomaz Mohoric, Asmaa Ali, Jeff Wiseman (Edelweiss Connect)

including collaboration with:

- Natalja Kurbatova and **Zifo** team
- Tom Luechtefeld and **InSilica** team
- Erik Schultes and **GOFAIR** team
- Collaborating partners on **RISK-HUNT3R**, **ONTOX**, **PrecisionTox** and **SSbD4CheM**
- Industry Case Study Partners

ECETOC Workshop
16-17 October 2024
Sophia Antipolis, France
Integrating AI into Chemical Safety Assessment

Overview

Edelweiss Connect

01 Issues and Methodology

Goals

- Develop robust evidence-based risk assessment practices and solutions
- Prepare reliable data and knowledge supporting risk assessment
- Use and refine AI approaches to enable knowledge management
- Optimise knowledge-driven AI approaches to assist risk assessment
- Support Human-AI collaboration including curation and analysis
- Test methods and practices against use cases and case studies
- Include open standards and best practices as a goal accelerator
- Align with developing AI practices and standards (trust, explainability, transparency, ethics, legal requirements)

Issues

- Significant variation in quality of data, *a lot of historical data is not very FAIR, may not have followed best practice protocols e.g., in kinetics modelling for dose range setting or has incomplete protocol and metadata description*
- Significant variation in quality of literature, *following a systematic review one may reject 95% or more of the reviewed articles as not matching evidence quality criteria*
- Significant amount of traditional high quality toxicology data is firewalled; *need for data sharing mechanisms which currently are limited*
- Preparation of FAIR data and knowledge is time and resource expensive; *quality has a cost but also needs convincing business cases*
- New approach methods supporting risk assessment often still have limited datasets and regulatory acceptance
- We have often agreed on the benefits of open standards and ontologies, *but coordinated support and engagement has often been disjointed with insufficient sustainable adoption*
- Risk assessment lacks a structured practice, *with significant variation in results obtained*
- New forms of AI often have limited expertise and thematic training, *provide results of uncertain quality*
- Increasing amount of “AI junk” on the internet and accelerated by social media lacking facts or truth, further exacerbating poor quality of digested *knowledge, curation of knowledge is becoming a a critical function*

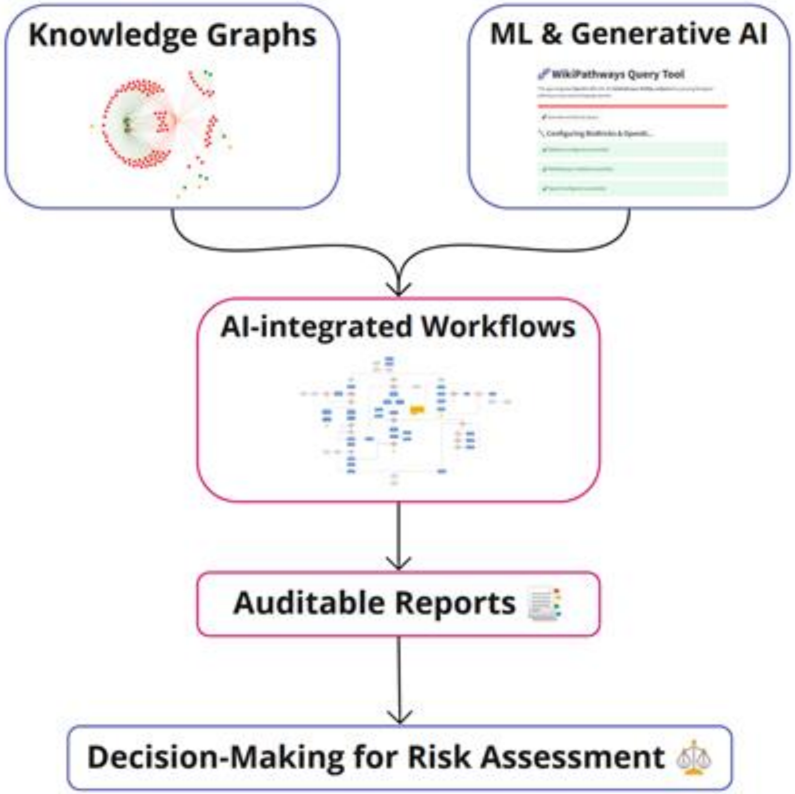
Approach

- Define Use Cases supporting risk assessment (e.g., search for all current information, define compound testing set, evaluate a concern related to metabolism, execute a task in a step of an SSbD workflow)
- Define (Draft) Business Cases (and test and refine them!)
- Agree problem formulations and case studies to be tested against
- Test use cases with and without use of (non-FAIR and FAIR-based) AI and compare results for quality, completeness, accuracy, transparency, costs, *including results from different sources of knowledge, and filtering/curation mechanisms*
- *Follow Principles of the OpenTox Framework 2.0 (original framework updated for developments over the last decade and especially with attention paid to recent AGI developments such as LLMs; article in preparation for OpenTox 2024 (25-29 November)* and In Vitro Toxicology)*

* https://www.opentox.net/events/virtual-conference-2024

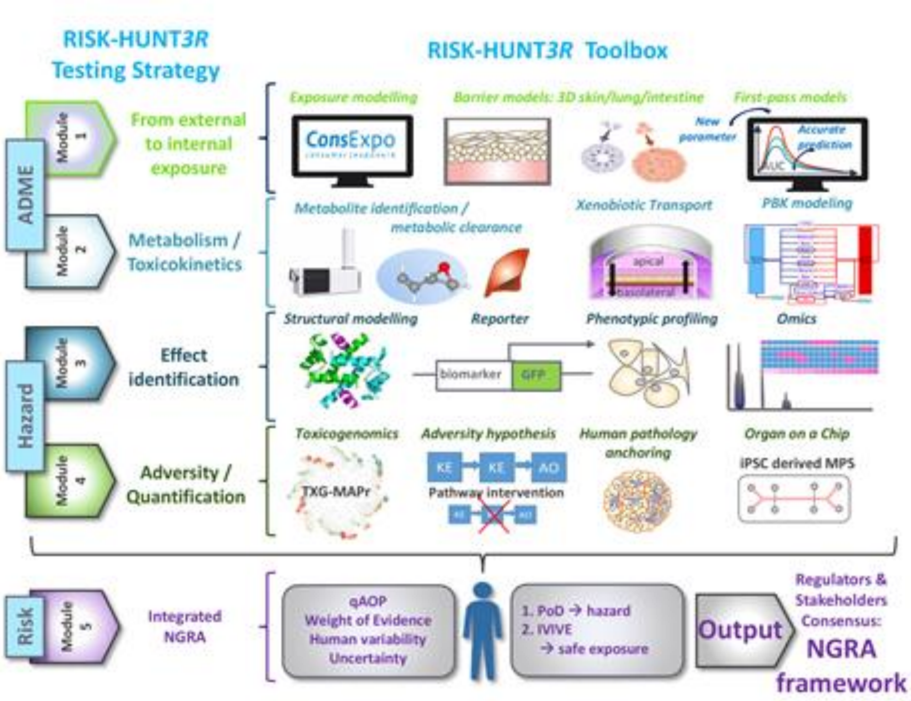
Risk Assessment Test Cases

1. Efficient retrieval of well-referenced knowledge via a knowledge graph.
2. Enhanced toxicology learning with ML and Generative AI.
3. AI-integrated workflows for risk assessment and decision-making.
4. Auditable reports with supporting data for risk assessments.

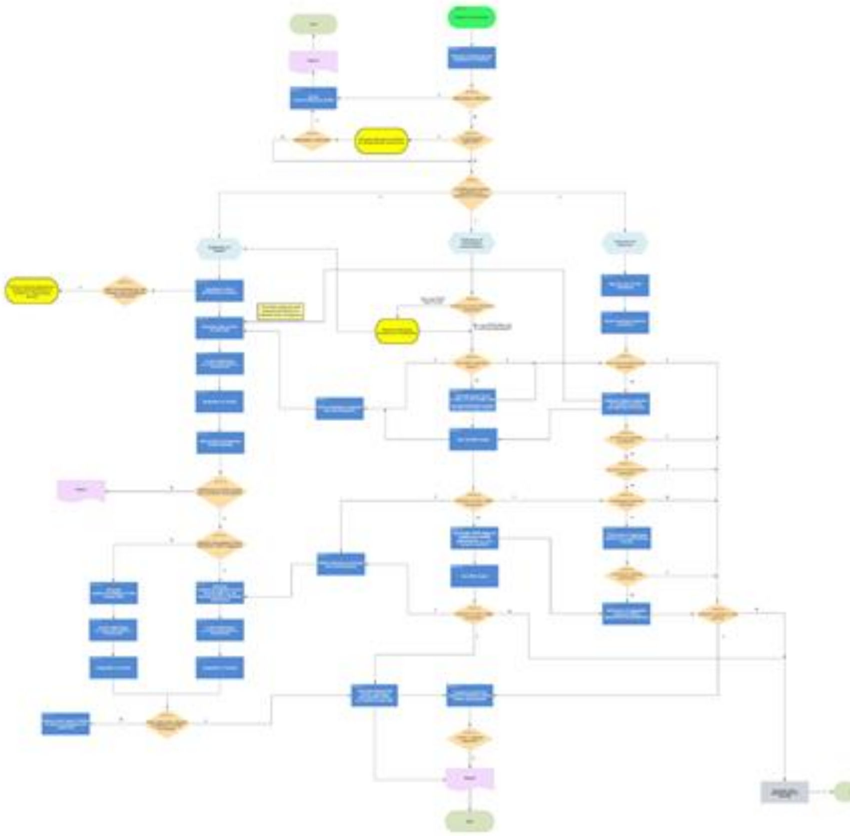


02 Cases and Applications

NGRA and ASPA Workflows

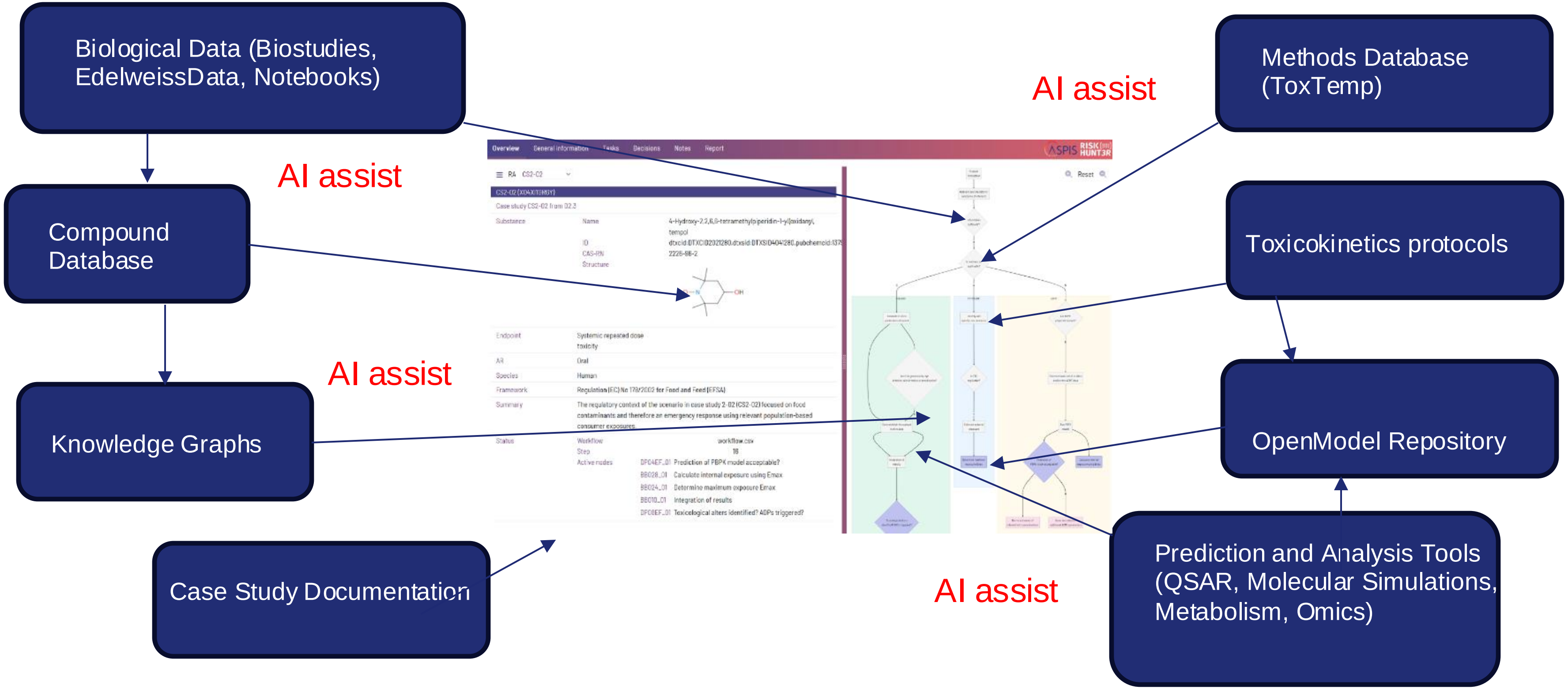


ASPA Workflow v2.0



https://risk-hunt3r.net/

Knowledge Infrastructure Support of ASPA Workflow



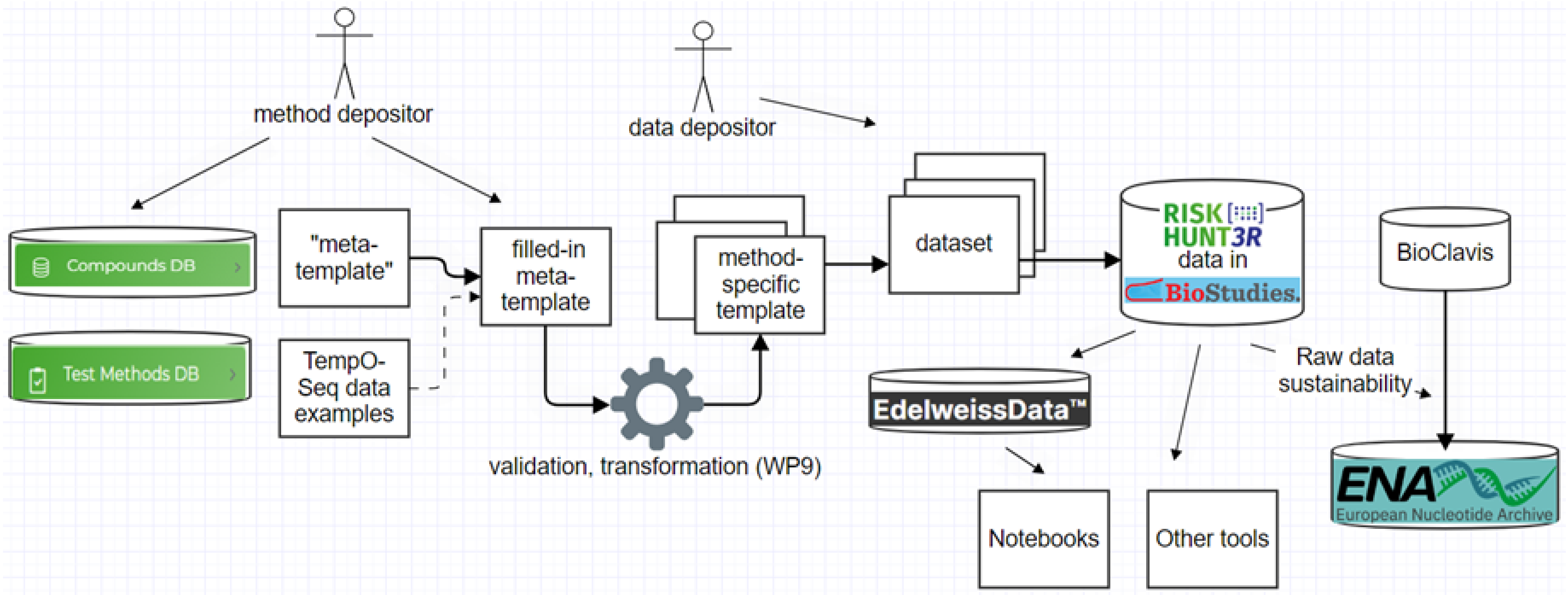
https://risk-hunt3r.net/

Safe and Sustainable by Design Infrastructure and Case Studies (SSbD4CheM)

https://www.ssbd4chem.eu/ (2024 - 2027)

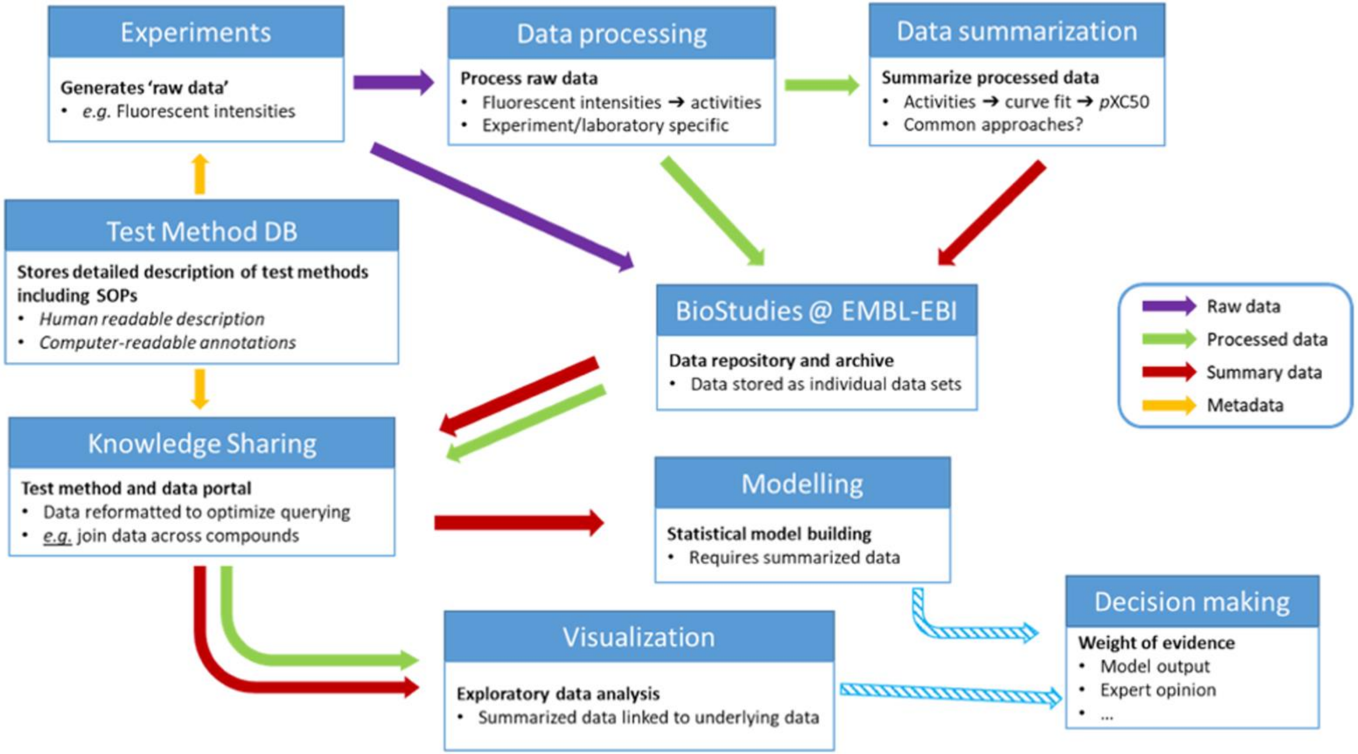
03 Sources of Knowledge

FAIR Data Processing Workflow



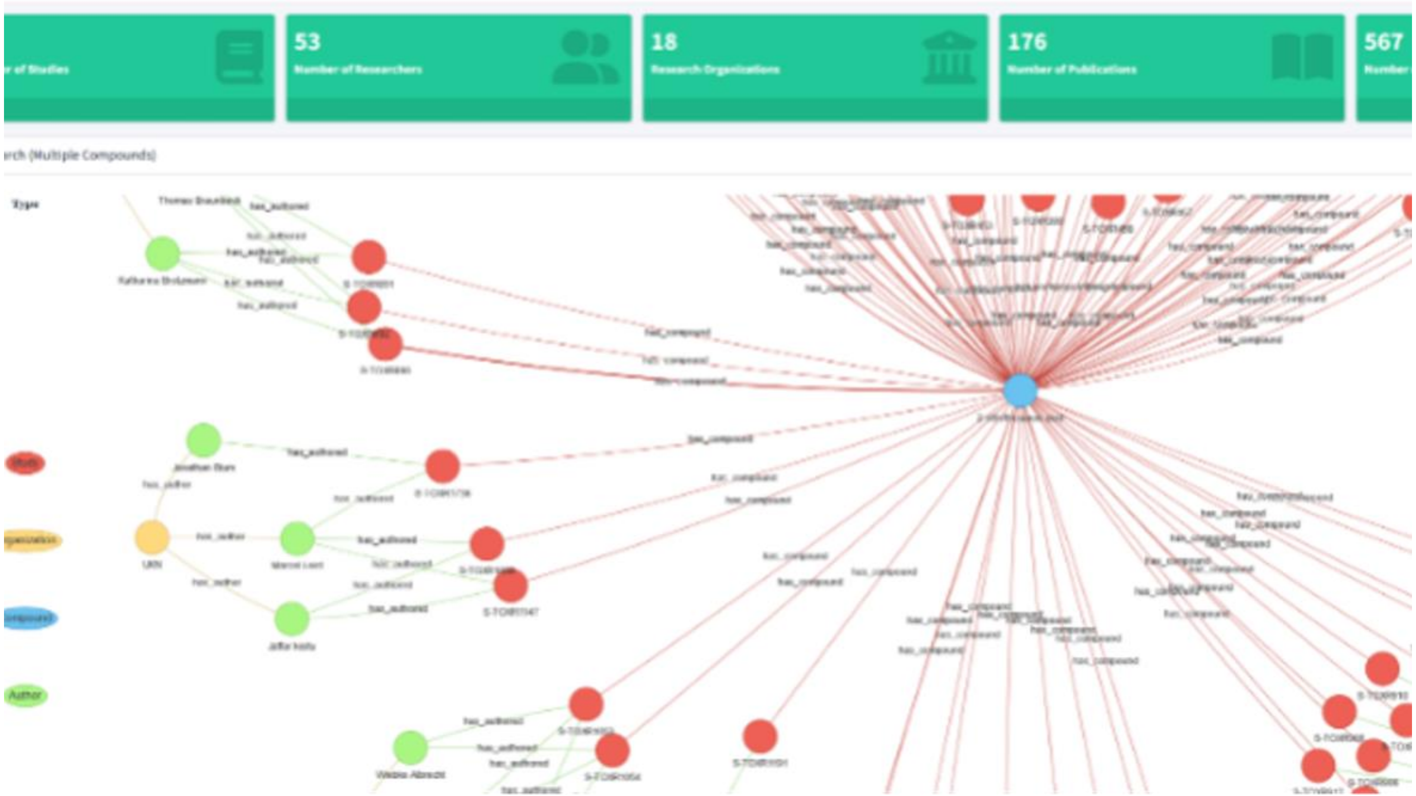
https://risk-hunt3r.net/

FAIR Knowledge Infrastructure (EU - ToxRisk)



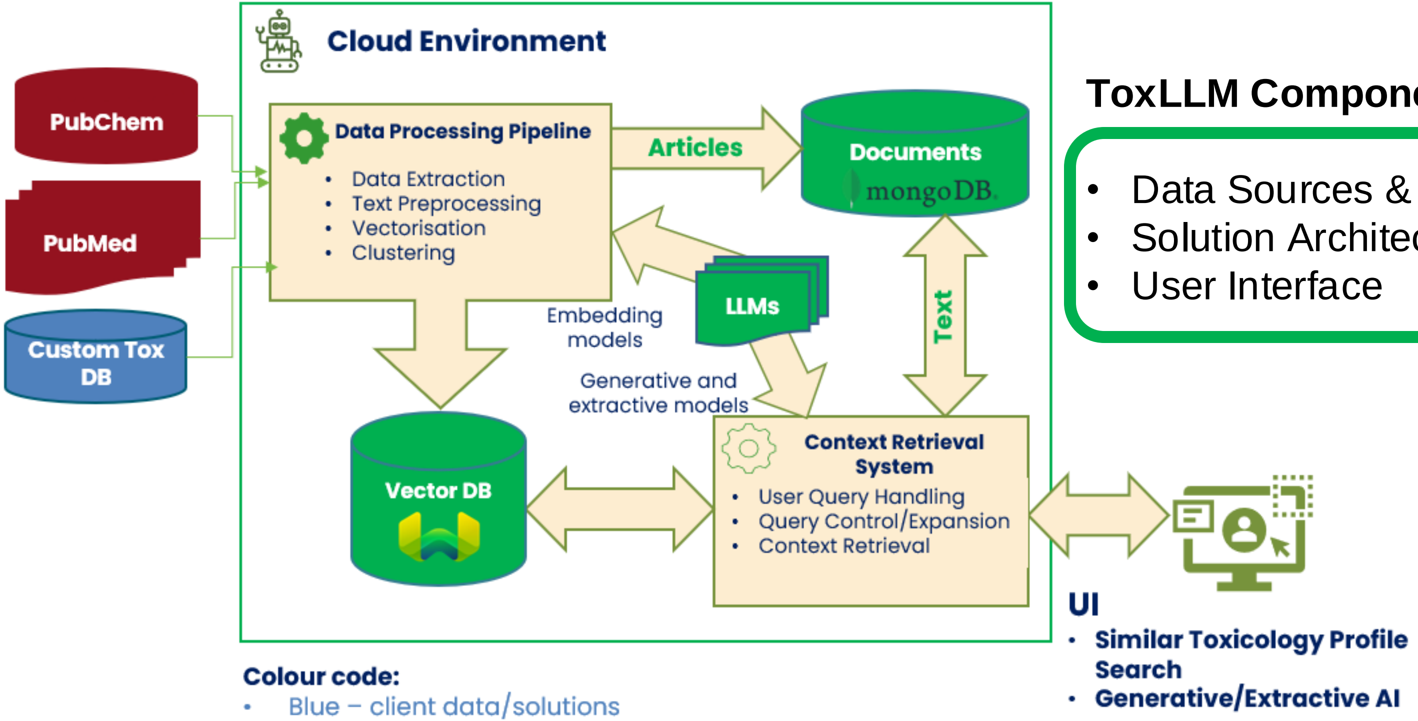
https://www.sciencedirect.com/science/article/pii/S0887233324001334 (2024)

FAIR Knowledge Infrastructure (EU-ToxRisk example)



https://risk-hunt3r.net/

Literature (ToxLLM Workflow)



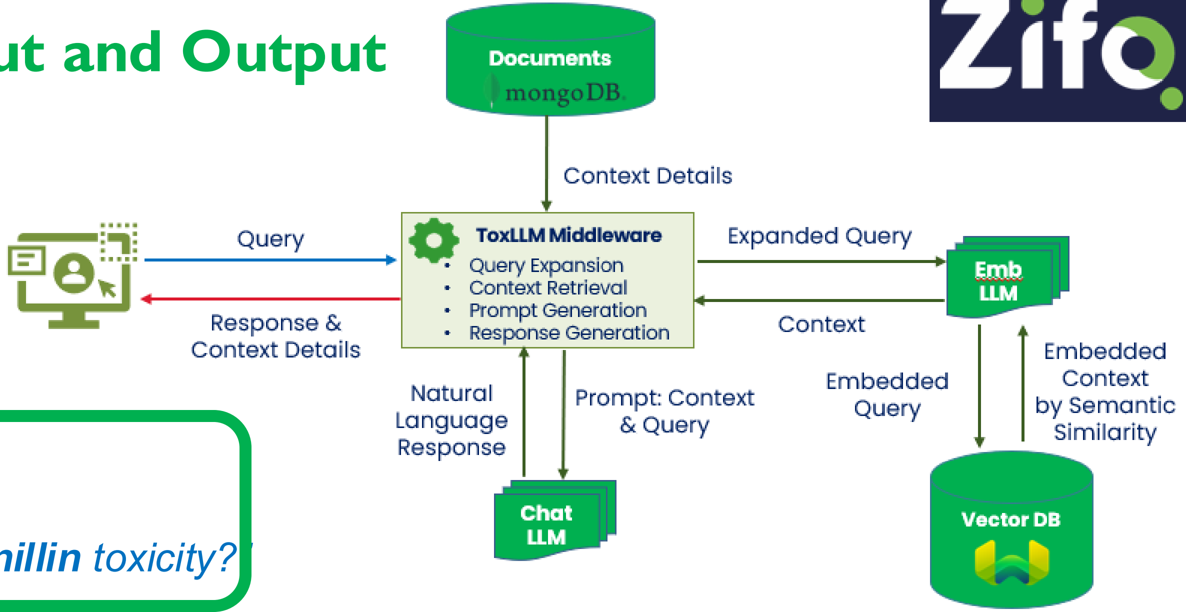
Colour code:

- Blue – client data/solutions
- Green – Zifo solutions
- Red - Public domain / Open-Source solutions

ToxLLM Components:

- Data Sources & Processing Pipelines
- Solution Architecture
- User Interface

Context Intelligent Input and Output



Input:

- Natural Language Query
  *"What do you know about **Vanillin** toxicity?"*

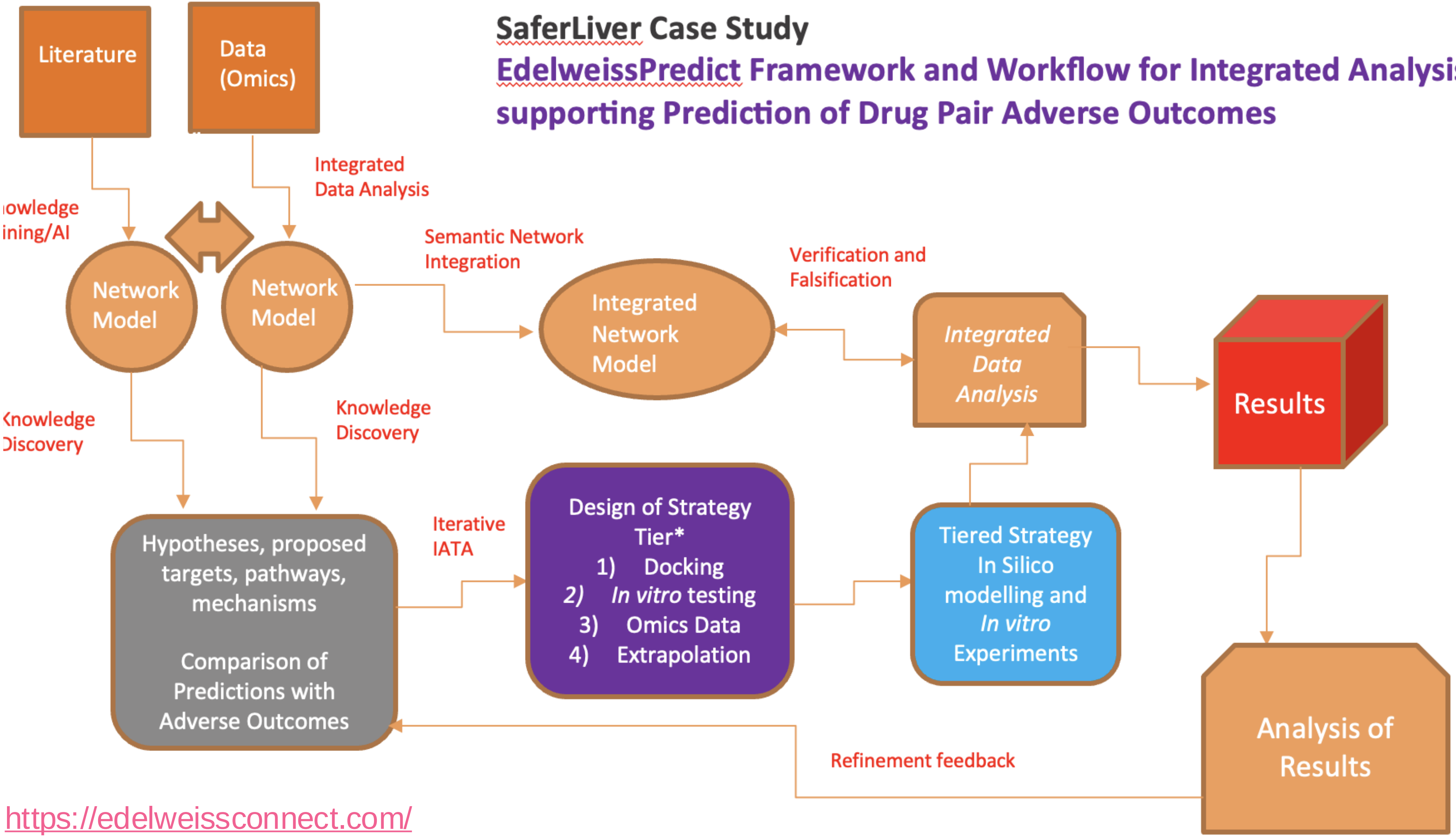
Output:

- Natural Language Response based on query and context plus context details (references and links)

"Based on the provided context, here is a concise answer summarizing every piece of information:
Vanillin is a widely used flavoring agent and fragrance ingredient that has been found to be non-toxic in most studies. However, high doses (>64 mg/kg body weight/day for 10 weeks) can cause growth depression and…"

SaferLiver Case Study

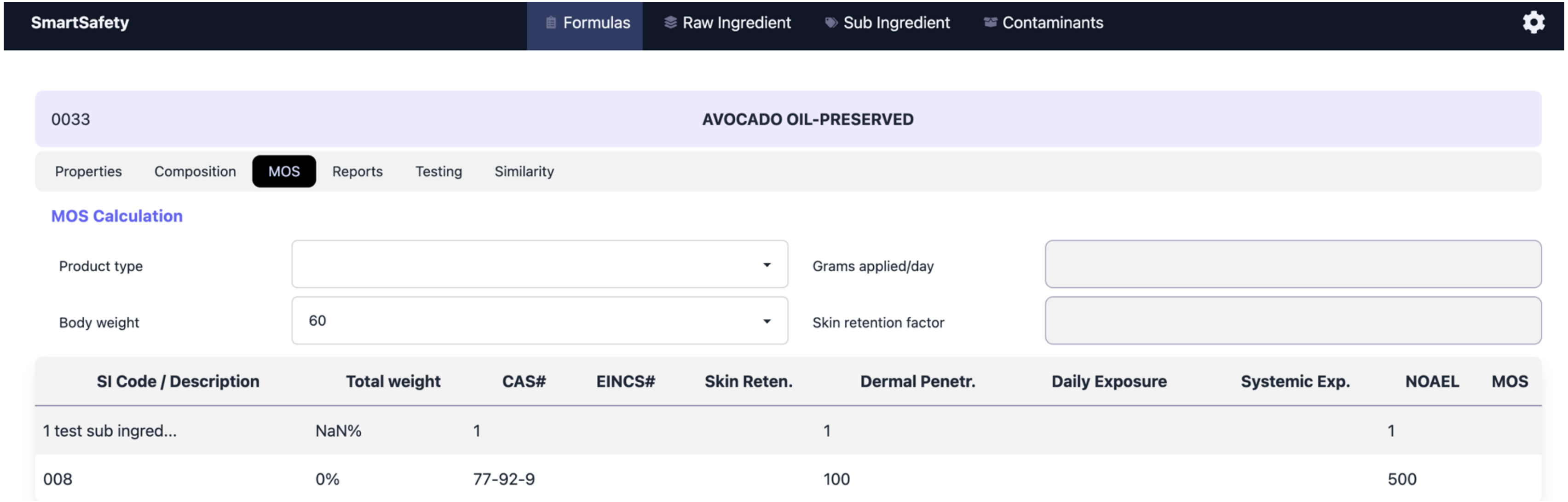
EdelweissPredict Framework and Workflow for Integrated Analysis supporting Prediction of Drug Pair Adverse Outcomes



https://edelweissconnect.com/

Product knowledge base (SmartSafety) – processing documents, generating reports in formulations and ingredients

AI assist on processing and extracting information from documents into ingredient and formulation profiles



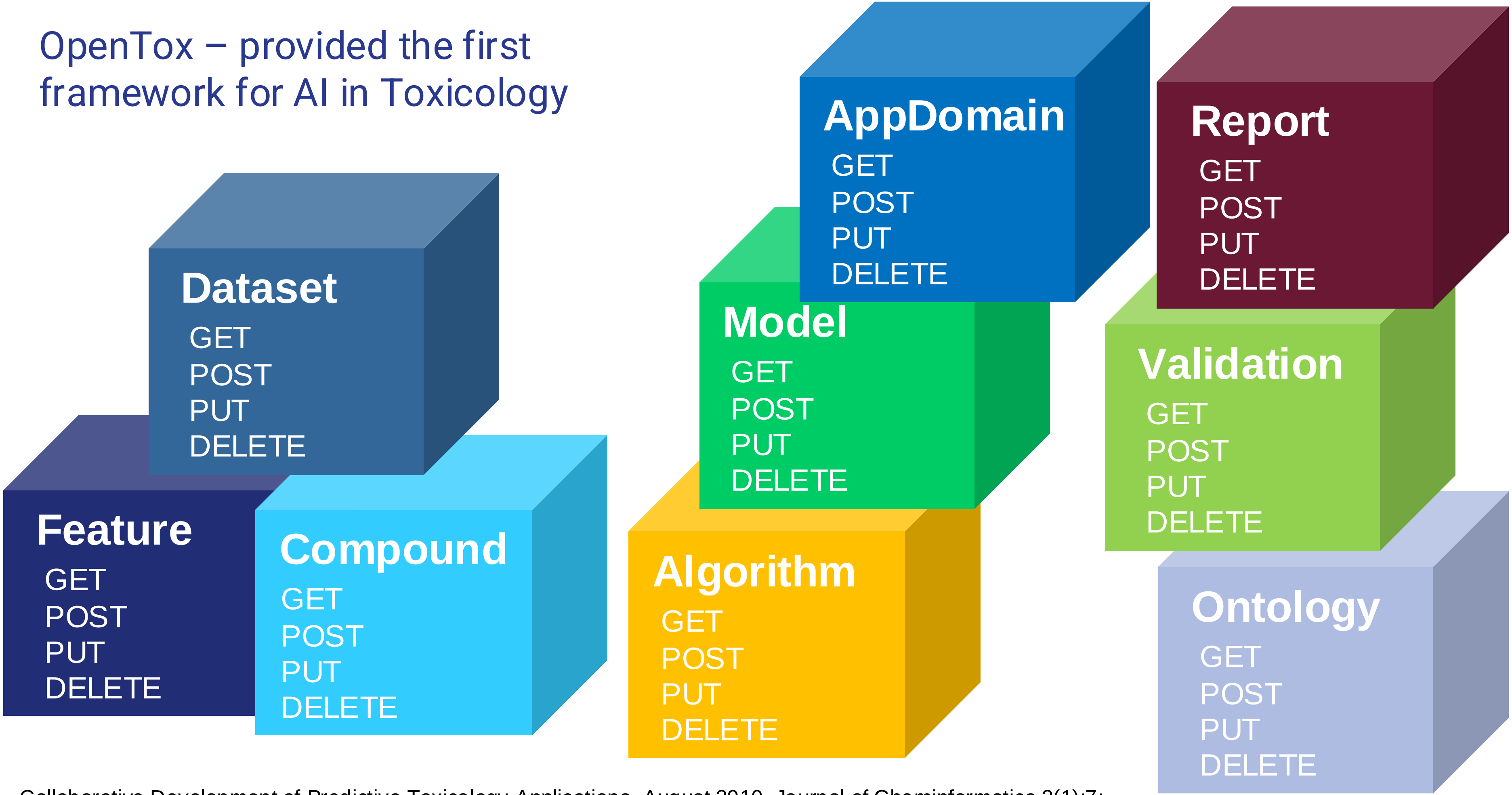
AI assist on generating reports (e.g., batches responding to a new or modified regulation)

Human – AI collaboration e.g, on curation, interpretation

https://smartsafety.edelweissconnect.com/

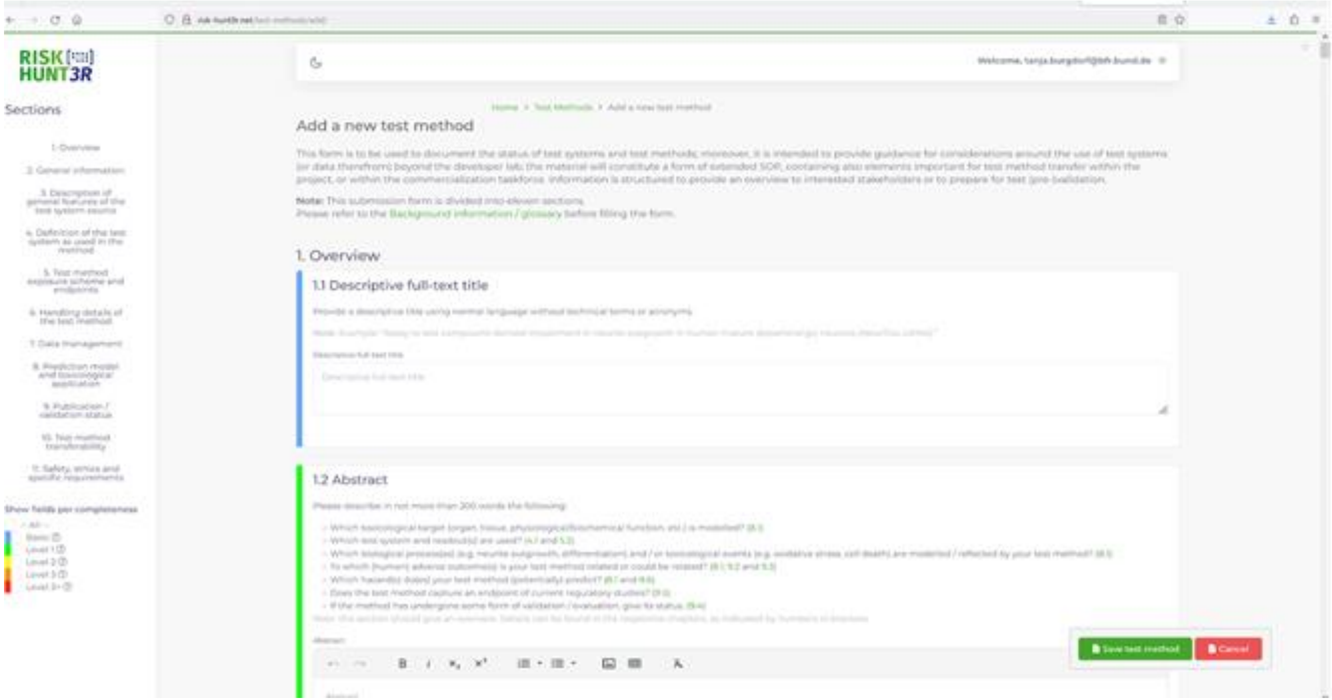
04 Practices and Standards

OpenTox – provided the first framework for AI in Toxicology



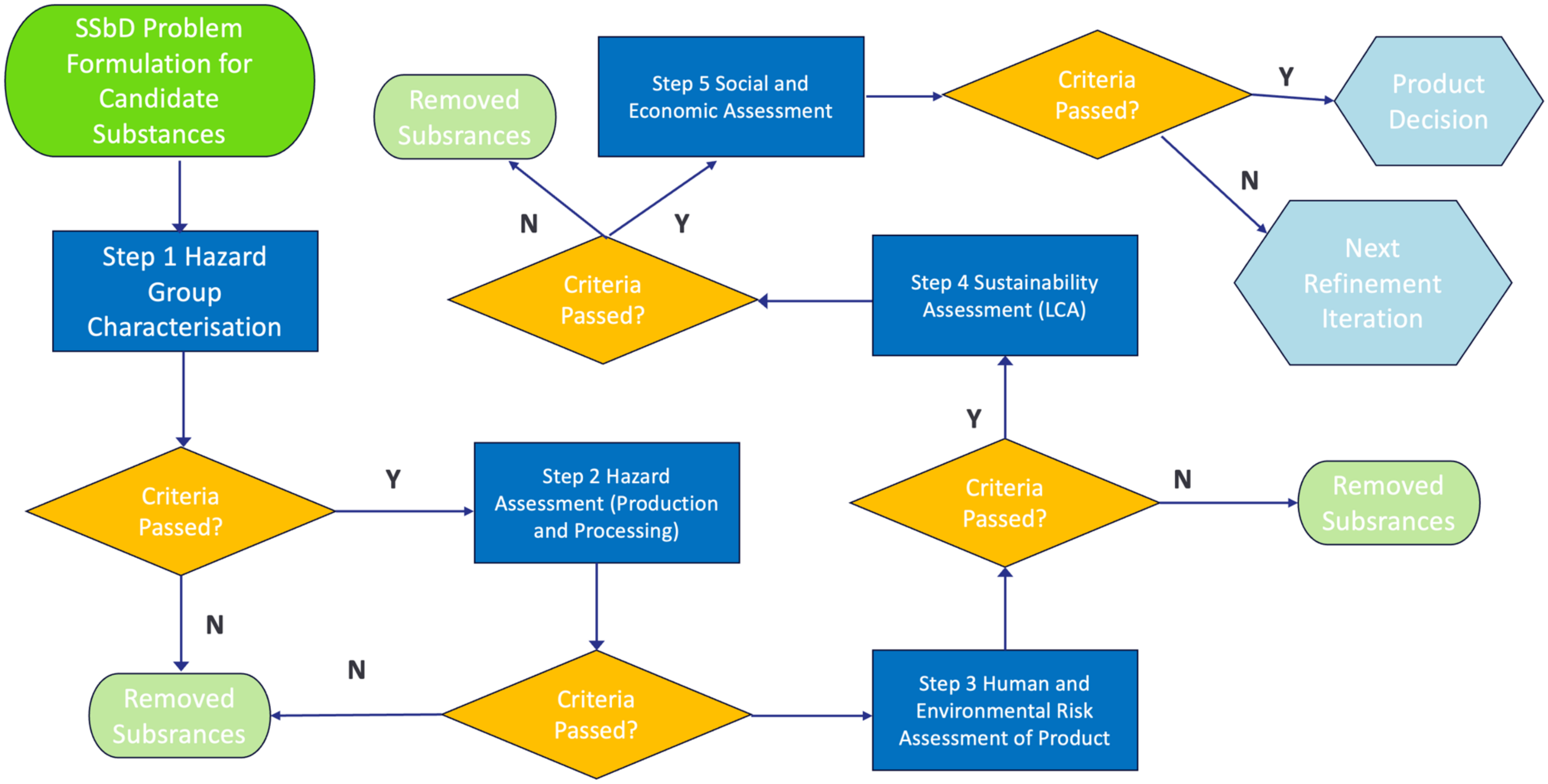
Collaborative Development of Predictive Toxicology Applications, August 2010, Journal of Cheminformatics 2(1):7; https://www.researchgate.net/publication/46107188_Collaborative_Development_of_Predictive_Toxicology_Applications

Method Documentation (ToxTemp)

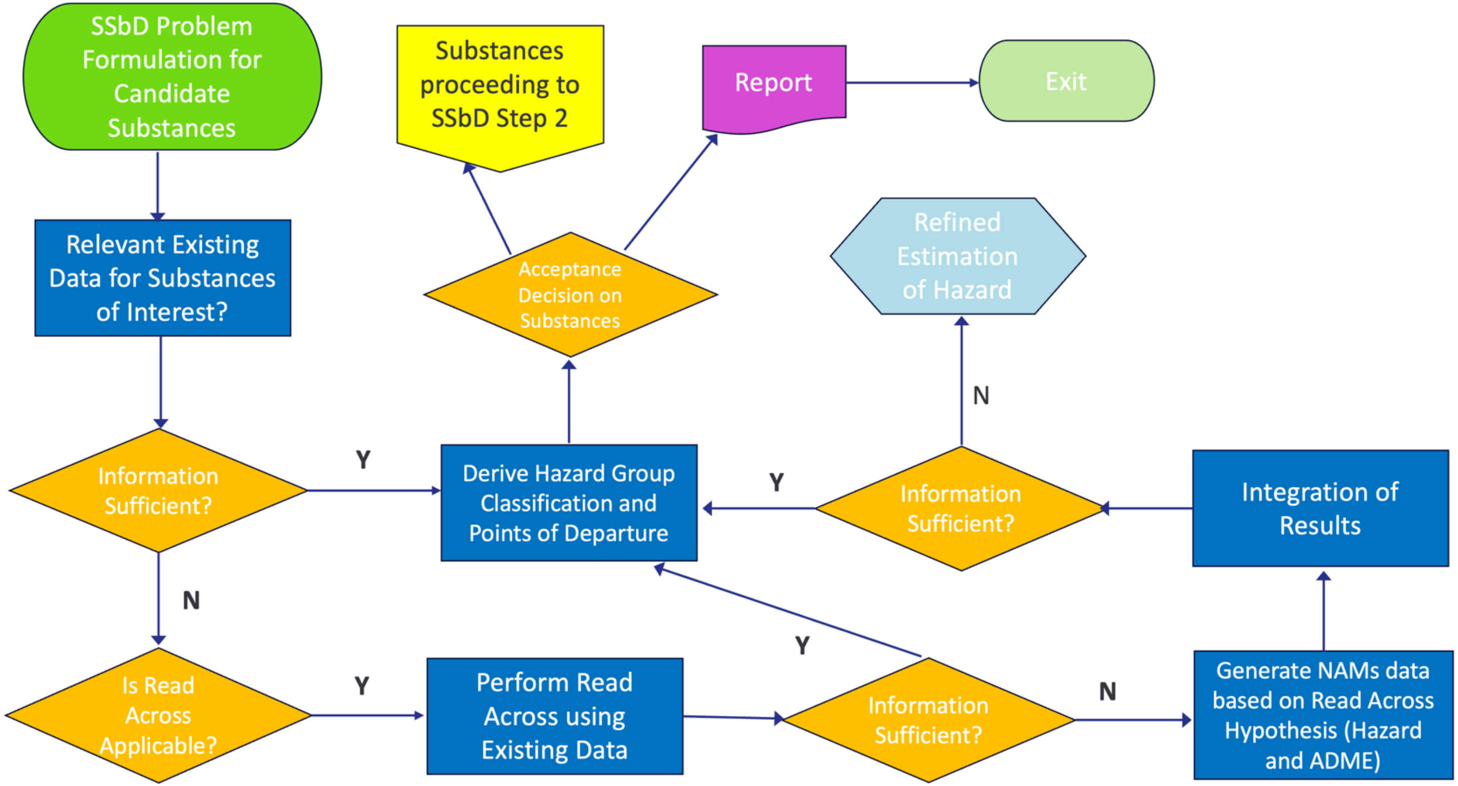


Template for the description of cell-based toxicological test methods to allow evaluation and regulatory use of the data, Alice Krebs et al (2019). https://www.altex.org/index.php/altex/article/view/1433

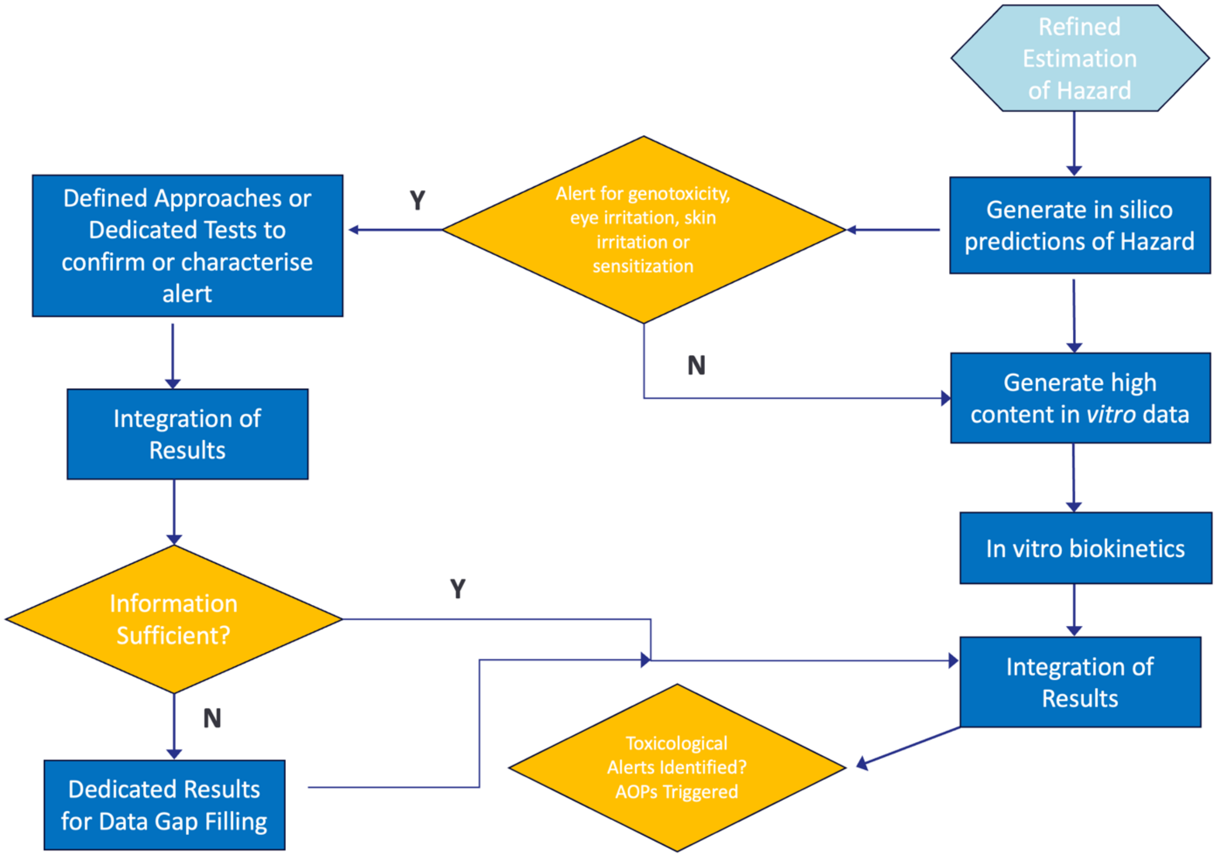
Workflows (Safer and Sustainable by Design (SSbD))



Workflows (SSbD Step 1)



Workflows (hazard refinement)



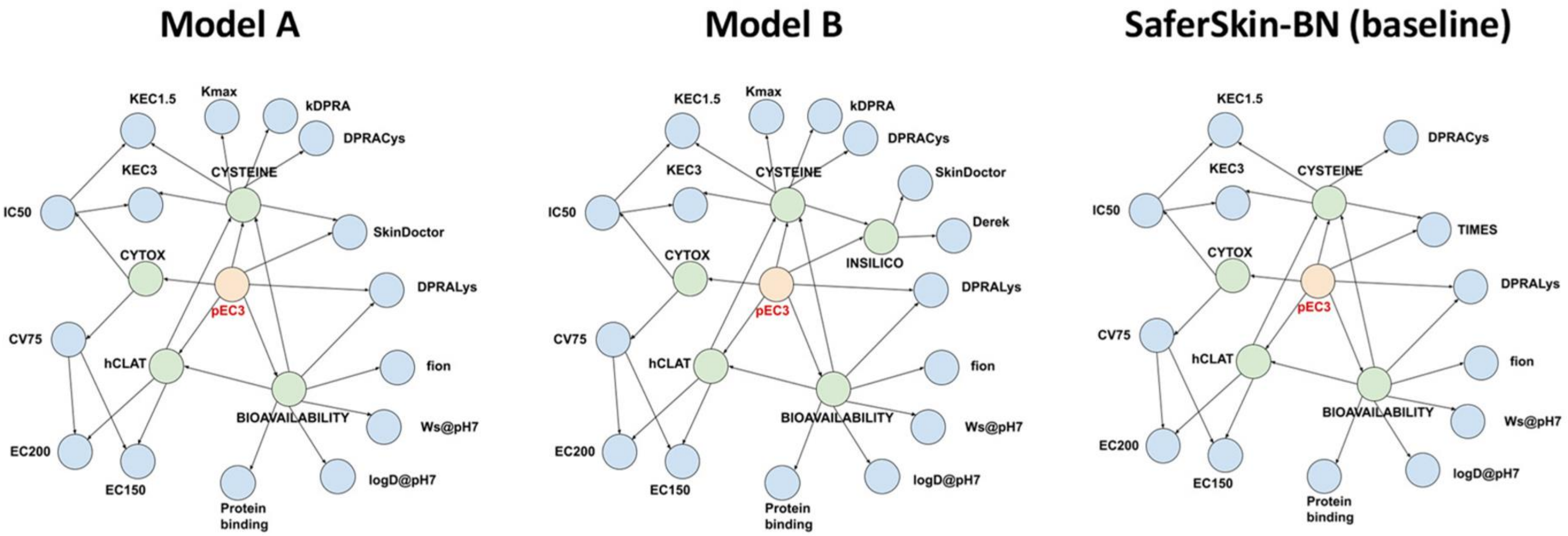
Incorporating data on FAIR Models into our NGRA Knowledge Infrastructure

- *In silico* models are used widely to support Next Generation Risk Assessment (NGRA).
- *In silico* models for NGRA must be findable and reproducible.
- FAIR principles have been applied to *in silico* models for chemical risk assessment.
- 18 FAIR principles for *in silico* models are proposed.
- FAIRification of *in silico* models is intended to increase use and acceptance.

Mark T.D. Cronin, Samuel J. Belfield, Katharine A. Briggs, Steven J. Enoch, James W. Firman, Markus Frericks, Clare Garrard, Peter H. Maccallum, Judith C. Madden, Manuel Pastor, Ferran Sanz, Inari Soininen, Despoina Sousoni

Making in silico predictive models for toxicology FAIR, Regulatory Toxicology and Pharmacology, Volume 140, 2023, 105385, ISSN 0273-2300, https://doi.org/10.1016/j.yrtph.2023.105385. (https://www.sciencedirect.com/science/article/pii/S0273230023000533)

Weight of Evidence (Bayesian Networks)



SaferSkin example implementing Defined Approach with Beiersdorf (2024)
https://www.mdpi.com/2305-6304/12/9/666

05 Legal and Ethics

Alignment with Legislation, DESIGN FOR AI-SAFETY (CBRN)

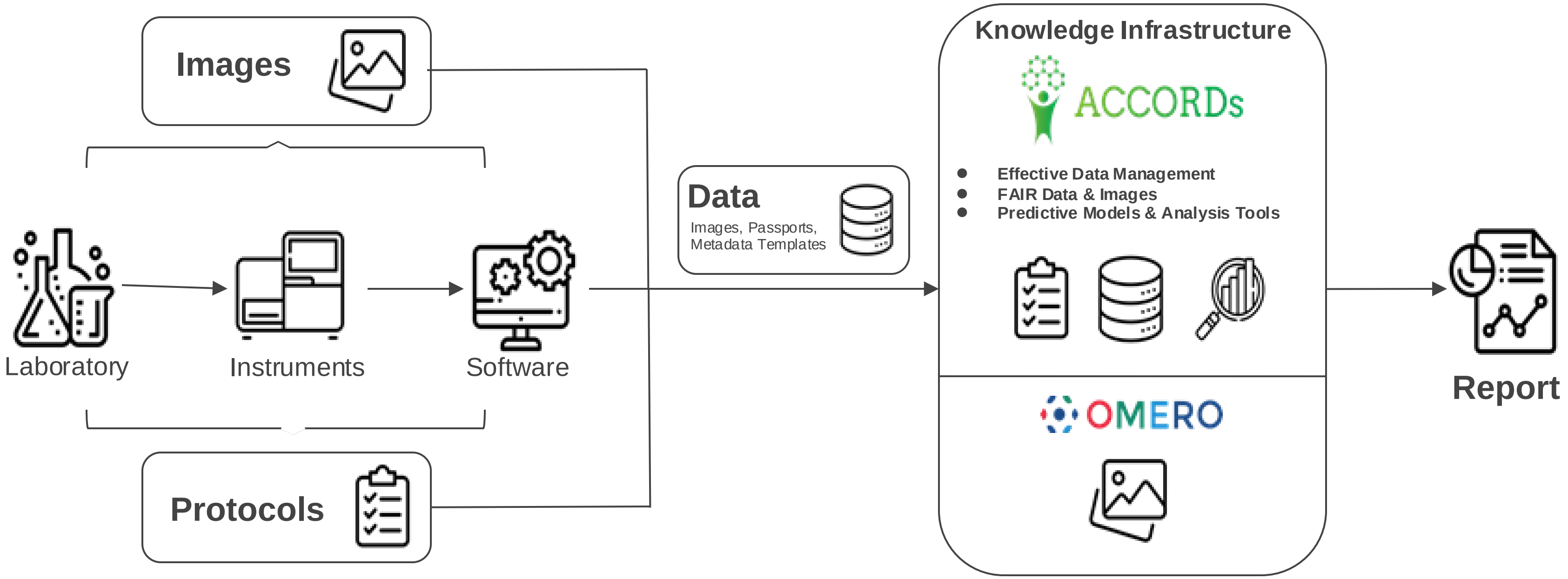
- Comply with AI safety regulations
  - NIST AI Risk Management Framework
  - EU AI Act
- Prevent misuse in context of CBRN threats (Chemical, Biological, Radiological, and Nuclear).
- We analyse how AI safety measures impact our use cases.

OpenAI Example

https://openai.com/safety

06 (FAIR) Use Case Example

Characterisation workflow



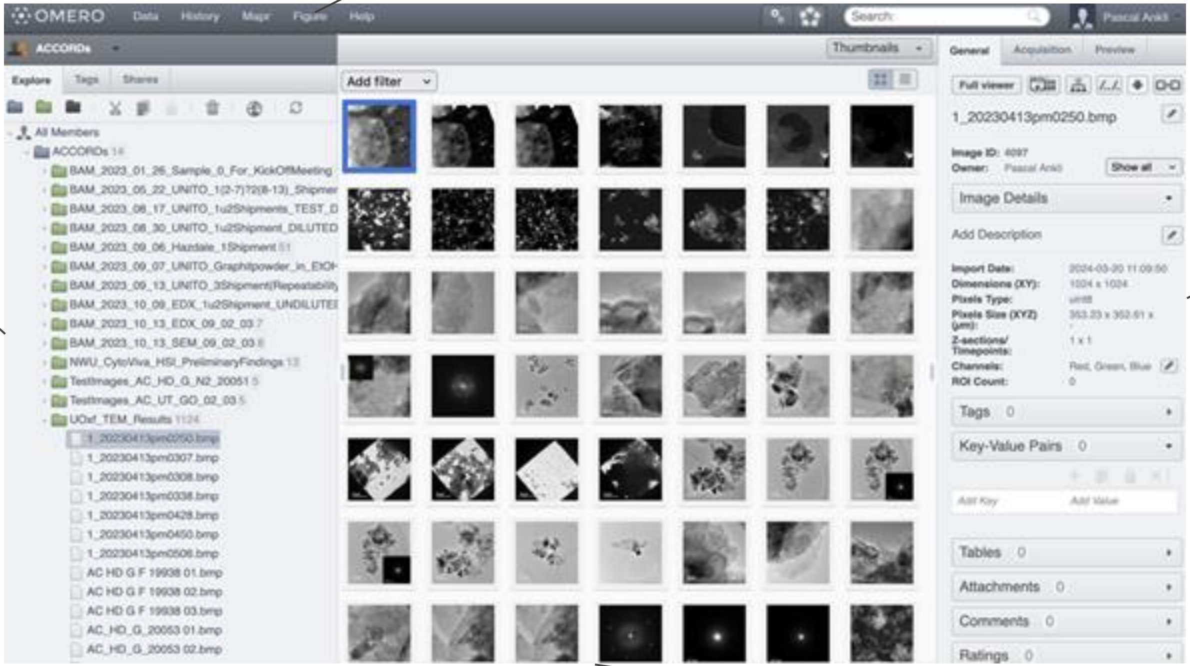
Knowledge infrastructure supporting image-based nanomaterial characterisation

https://accordsproject.com/

Image related data in ACCORDs Image Repository (OMERO)

Easy figure creation
for publication

Organised folder
structure



Data and metadata
overview

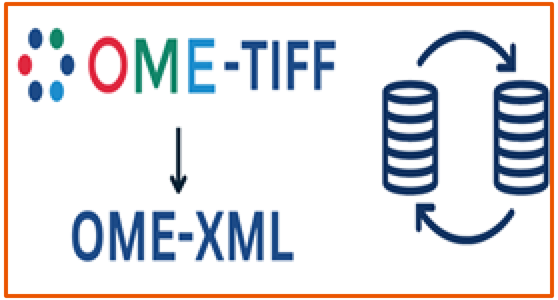
Analysis, Editing etc.

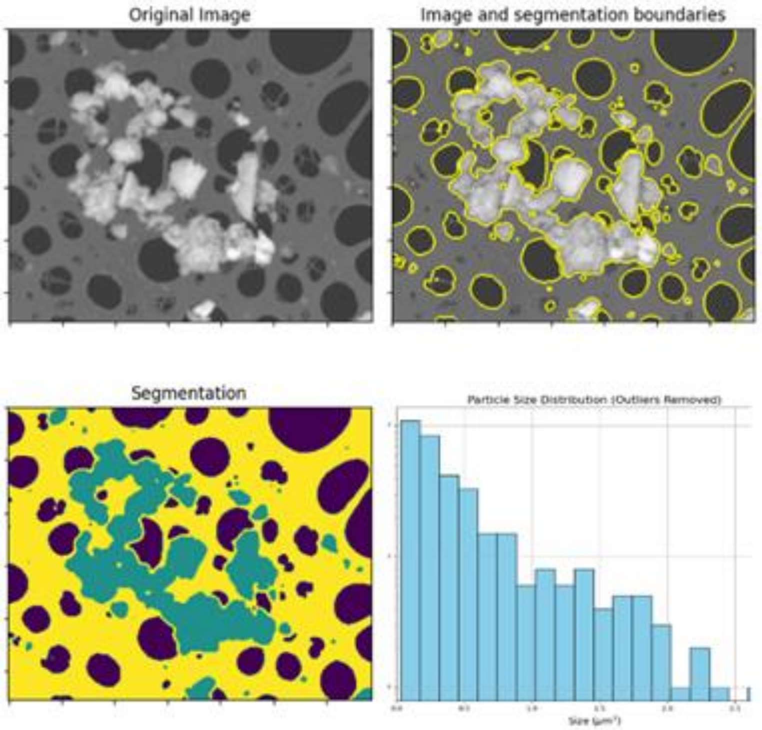
https://accordsproject.com/

Automated Metadata Extraction, Image Annotation and Analysis

OME-Formats for FAIR Images

With image conversation to OME-TIFF and OME-XML we can annotate customised metadata to images and image data sets in an open format supporting FAIR goals

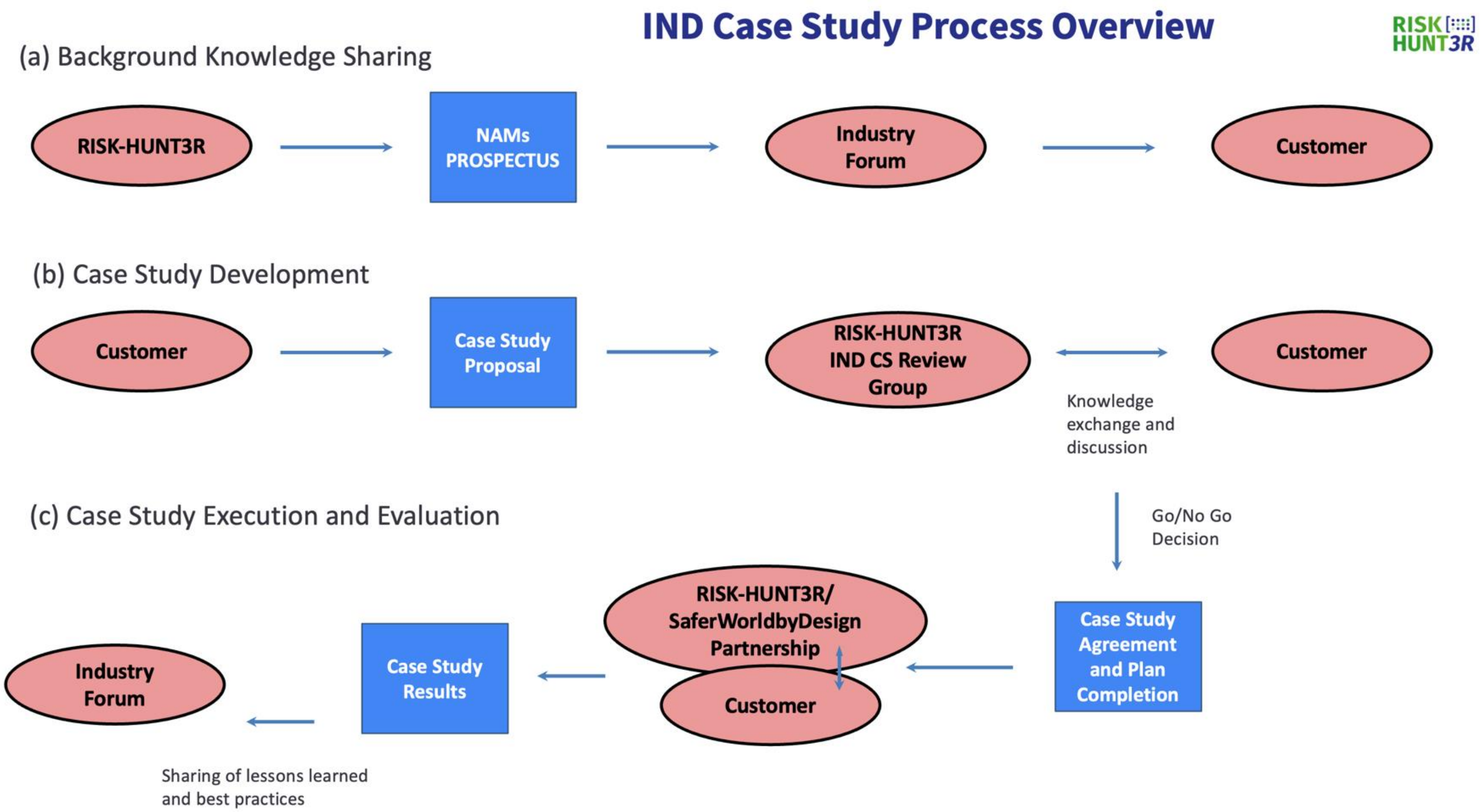




https://accordsproject.com/

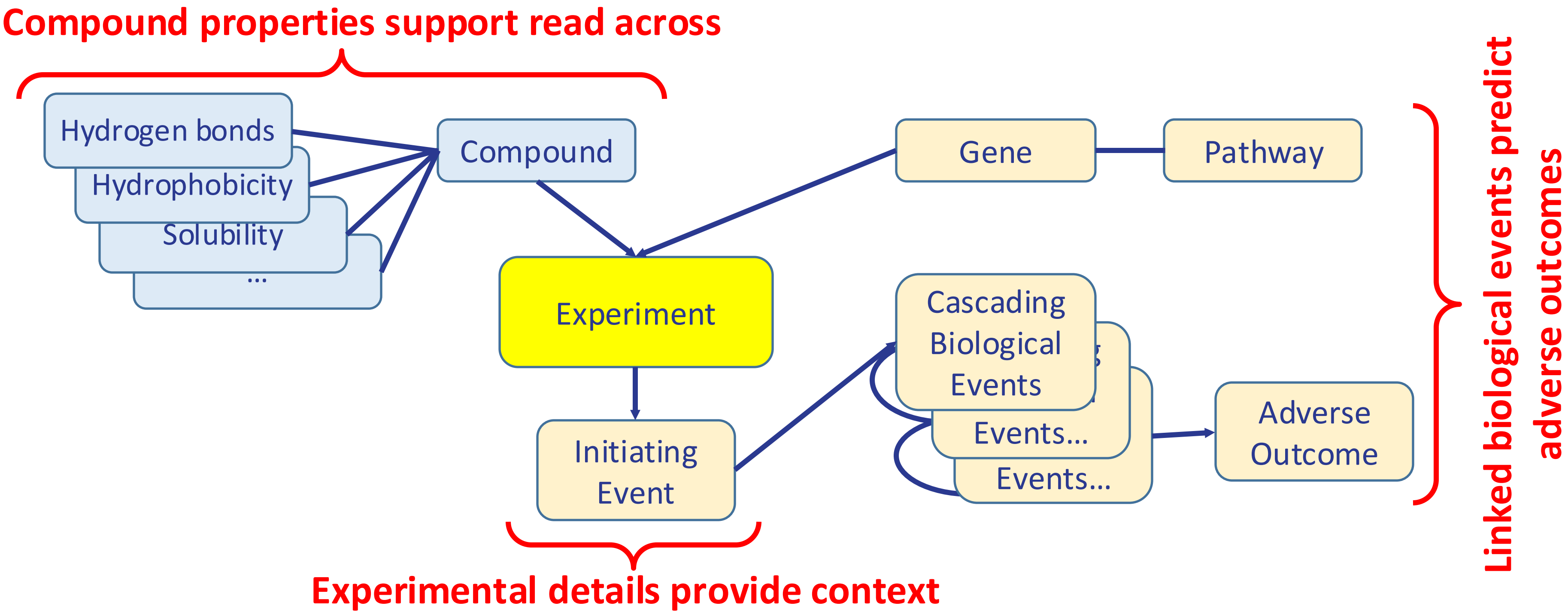
07 Outlook

Tackling Challenges

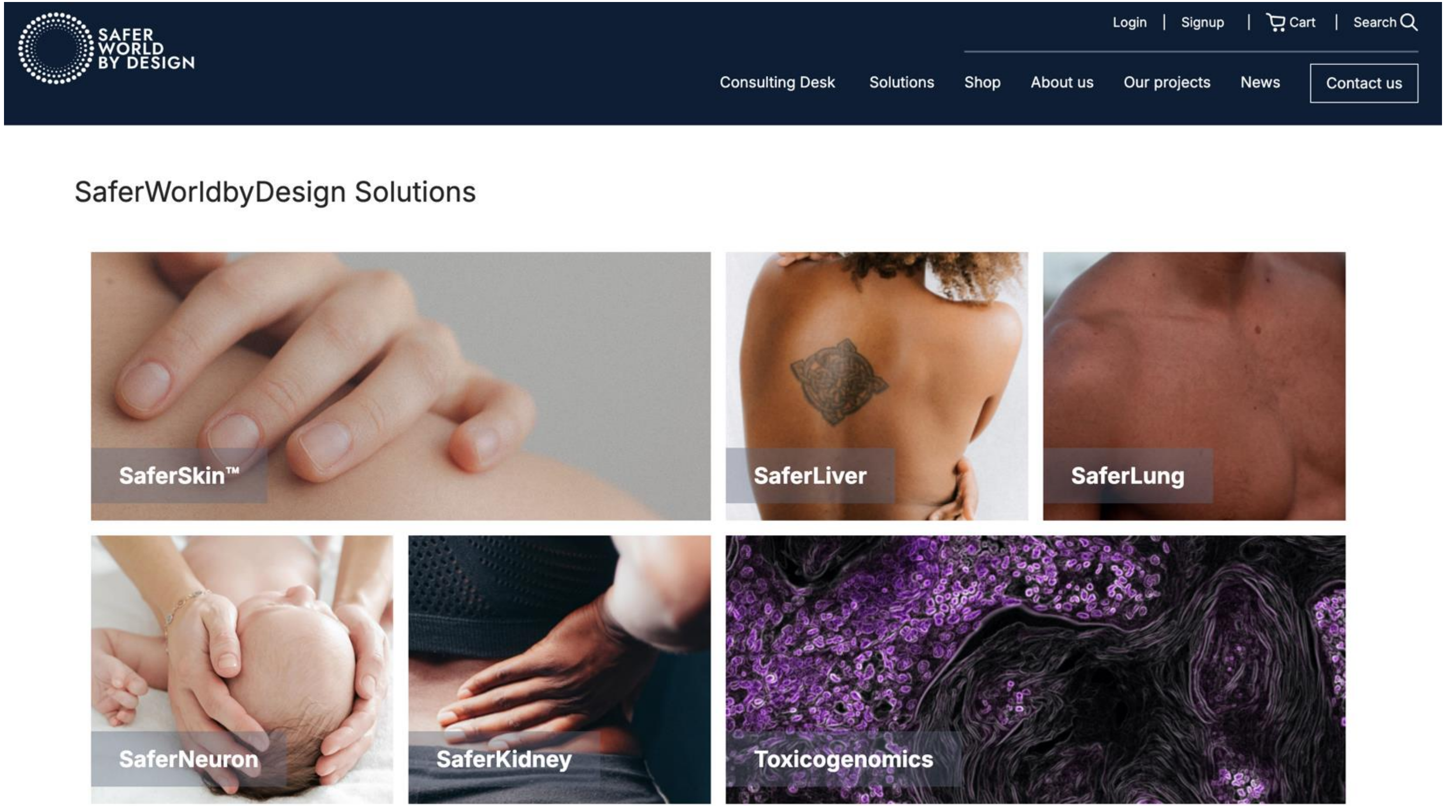


Contact: Barry.Hardy@edelweissconnect.com

Integration of Chemical and Biological Information, Testing of Models, Knowledge Graphs and AI-Assisted workflows



Integrated Solution Development



https://saferworldbydesign.com/

We strive for collaborative innovation working with stakeholders!

At Edelweiss Connect, we foster creativity, collaboration, and curiosity to drive continuous improvement in risk assessment and data management. By maintaining an agile and adaptive mindset, we navigate challenges, refine ideas, and explore new frontiers to provide solutions in safe design, risk assessment and sustainability.

https://www.edelweissconnect.com/

Acknowledgements

The ACCORDs and SSbD4CheM projects receive funding from the European Union's Horizon Europe Research & Innovation Programme under grant agreements no. 101092796 and 101138475 respectively. Associated Partners (i.e. (a) Swiss Partners and (b) UK Partners) have received national funding from (a) the Swiss State Secretariat for Education, Research and Innovation (SERI), and (b) Innovate UK.

The RISK-HUNT3R project receives funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No. 964537 and is part of the ASPIS cluster also consisting of EU-funded projects ONTOX and PrecisionTox.

Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Health and Digital Executive Agency (HADEA). Neither the European Union nor the granting authority can be held responsible for them.