







Report of the First ONTOX Stakeholder Network Meeting: Digging Under the Surface of ONTOX Together With the Stakeholders

Alternatives to Laboratory Animals
2024, Vol. 52(2) 117–131
© The Author(s) 2024



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/02611929231225730
journals.sagepub.com/home/atl



Michael G. Diemar¹, Mathieu Vinken², Marc Teunis³, Cyrille A.M. Krul³, Francois Busquet⁴, Julia Dominika Zajac², Helena Kandarova⁵, Raffaella Corvi⁶, Matteo Z. Rosso⁷, Anastasiia Kharina⁷, Louise Stab Bryndum⁸, Michael Santillo⁹ , Denise Bloch¹⁰ , Olena Kucheryavenko¹¹ , Demosthenes Panagiotakos¹², Vera Rogiers^{2,13}, Manon Beekhuijzen¹⁴, Arianna Giusti¹⁵, Abdulkarim Najjar^{16,17} , Carol Courage¹⁷, Torben Koenig¹⁷, Susanne Kolle¹⁸, Harrie Boonen¹⁹, Stephane Dhalluin²⁰, Julie Boberg²¹, Boris P. Müller²², Predrag Kukic²³ , Merel Ritskes-Hoitinga^{24,25} , Elena Grasselli^{26,27}, Tamara Zietek²⁸, Gilly Stoddart²⁹, Harm J. Heusinkveld³⁰, Jose V. Castell³¹, Emilio Benfenati³², Huan Yang³³, Simón Perera^{34,35}, Alicia Pains³³, Nynke I. Kramer³⁶, Thomas Hartung³⁷, Manoe Janssen³⁸, Ellen Fritsche^{39,40,41}, Danyel G.J. Jennen⁴², Matteo Piumatti⁴, James Rathman⁴³, Jörg Marusczyk⁴³, Lucia Milec⁵ and Erwin L. Roggen¹

¹3Rs Management and Consulting ApS, Kongens Lyngby, Denmark

²Department of Pharmaceutical and Pharmacological Sciences, Vrije Universiteit Brussel, Brussel, Belgium

³Innovative Testing in Life Sciences & Chemistry, University of Applied Sciences, Utrecht, The Netherlands

⁴Altortox, Ixelles, Belgium

⁵Institute of Experimental Pharmacology and Toxicology, Centre of Experimental Medicine, Slovak Academy of Sciences, Bratislava, Slovakia

⁶European Commission, Joint Research Centre (JRC), Ispra, Italy

⁷Angel Consulting S.a.S., Genova, Italy

⁸Danish Environmental Protection Agency, Copenhagen, Denmark

⁹Center for Food Safety and Applied Nutrition (CFSAN), US Food and Drug Administration (FDA), Laurel, MD, USA

¹⁰Department of Pesticides Safety, German Federal Institute for Risk Assessment (BfR), Berlin, Germany

¹¹Department of Chemicals and Product Safety, German Federal Institute for Risk Assessment (BfR), Berlin, Germany

¹²SCHEER/EC–Scientific Committee on Health, Environmental and Emerging Risks/European Commission, European Commission, Brussels, Belgium

¹³SCCS–Scientific Committee on Consumer Safety, European Commission, Brussels, Belgium

¹⁴Charles River Laboratories, 's-Hertogenbosch, The Netherlands

¹⁵Cosmetics Europe–The Personal Care Association, Brussels, Belgium

¹⁶Beiersdorf AG, Hamburg, Germany

¹⁷International Collaboration on Cosmetics Safety (ICCS), Mt Royal, NJ, USA

¹⁸Experimental Toxicology and Ecology, BASF SE, Ludwigshafen am Rhein, Germany (for CEFIC)

¹⁹H. Lundbeck, Valby, Denmark

²⁰L'Oreal, Paris, France

²¹Novo Nordisk, Bagsværd, Denmark

²²Symrise AG, Holzminden, Germany

²³Safety & Environmental Assurance Centre, Unilever, Bedfordshire, UK

²⁴Institute for Risk Assessment Sciences Toxicology, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands

²⁵Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

²⁶Department of Earth, Environment, and Life Science, University of Genoa, Genova, Italy

²⁷Interuniversity Center for the Promotion of 3R Principles in Teaching and Research (Centro 3R), Pisa, Italy

²⁸Doctors Against Animal Experiments, Cologne, Germany

²⁹PETA Science Consortium International e.V., Stuttgart, Germany

Abstract

The first Stakeholder Network Meeting of the EU Horizon 2020-funded ONTOX project was held on 13–14 March 2023, in Brussels, Belgium. The discussion centred around identifying specific challenges, barriers and drivers in relation to the implementation of non-animal new approach methodologies (NAMs) and probabilistic risk assessment (PRA), in order to help address the issues and rank them according to their associated level of difficulty. ONTOX aims to advance the assessment of chemical risk to humans, without the use of animal testing, by developing non-animal NAMs and PRA in line with 21st century toxicity testing principles. Stakeholder groups (regulatory authorities, companies, academia, non-governmental organisations) were identified and invited to participate in a meeting and a survey, by which their current position in relation to the implementation of NAMs and PRA was ascertained, as well as specific challenges and drivers highlighted. The survey analysis revealed areas of agreement and disagreement among stakeholders on topics such as capacity building, sustainability, regulatory acceptance, validation of adverse outcome pathways, acceptance of artificial intelligence (AI) in risk assessment, and guaranteeing consumer safety. The stakeholder network meeting resulted in the identification of barriers, drivers and specific challenges that need to be addressed. Breakout groups discussed topics such as hazard versus risk assessment, future reliance on AI and machine learning, regulatory requirements for industry and sustainability of the ONTOX Hub platform. The outputs from these discussions provided insights for overcoming barriers and leveraging drivers for implementing NAMs and PRA. It was concluded that there is a continued need for stakeholder engagement, including the organisation of a ‘hackathon’ to tackle challenges, to ensure the successful implementation of NAMs and PRA in chemical risk assessment.

Keywords

artificial intelligence, confidence, end-user acceptance, hackathon, NAMs, new approach methodologies, non-animal, next generation risk assessment, probabilistic risk assessment, stakeholder network, toxicology

Introduction

The goal of the Horizon 2020 ONTOX project¹ — ontology-driven and artificial intelligence-based repeated dose toxicity testing of chemicals for next generation risk assessment (<https://www.ontox-project.eu>) — is to provide a generic functional and sustainable solution for advancing *the assessment of chemical risk to humans*, without the use of animals, in line with the 21st century toxicity testing and next-generation risk assessment (NGRA) principles. This solution consists of a strategy to create *new approach methodologies* (NAMs), in order to predict systemic repeated dose toxicity effects that, in

combination with tailored exposure assessment, will enable human risk assessment. The ONTOX definition of NAMs is as follows: an artificial intelligence-based strategy linked with a battery of *in vitro* assays and *in silico* tools for hazard prediction, to be combined with customised exposure assessment for the purpose of human risk assessment.

While the envisaged strategy is to be applicable to any type of chemical and systemic repeated dose toxicity effect, proof-of-concept will be established for six specific NAMs addressing adversities in the liver (steatosis and cholestasis), kidneys (tubular necrosis and crystallopathy) and the developing brain (neural tube closure and cognitive function

³⁰Dutch National Institute for Public Health and the Environment, Centre for Health Protection, Bilthoven, The Netherlands

³¹Department of Biochemistry, IIS Hospital La Fe and CIBERHED, University of Valencia, Valencia, Spain

³²Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Milano, Italy

³³esqLABS GmbH, Saterland, Germany

³⁴ProtoQSAR SL, Centro Europeo de Empresas Innovadoras, Paterna, Spain

³⁵Departament de Medicina i Ciències de la Vida, Institut de Biologia Evolutiva (CSIC-UPF), Universitat Pompeu Fabra, Barcelona, Spain

³⁶Toxicology Chair Group, Wageningen University, Wageningen, The Netherlands.

³⁷Center for Alternatives to Animal Testing Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA

³⁸The Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Utrecht, The Netherlands

³⁹IUF–Leibniz Research Institute for Environmental Medicine, Düsseldorf, Germany

⁴⁰Medical Faculty, Heinrich-Heine-University, Düsseldorf, Germany

⁴¹DNTOX GmbH, Düsseldorf, Germany

⁴²Department of Toxicogenomics, GROW School for Oncology and Reproduction, Maastricht University, Maastricht, The Netherlands

⁴³Molecular Networks GMBH Computerchemie, Nuremberg, Germany

Corresponding author:

Michael G. Diemar, 3Rs Management and Consulting ApS, Asavænget 14, 2800 Kongens Lyngby, Denmark.

Email: mgd@3rsmc.onmicrosoft.com

defects), that are induced by chemicals associated with pharmaceuticals, cosmetics, foods and biocides. These six NAMs will each consist of a computational system that incorporates cutting-edge artificial intelligence (AI), and will be based upon available biological/mechanistic, toxicological/epidemiological, physicochemical and kinetic data. The data will be consecutively integrated in physiological maps, quantitative adverse outcome pathway (qAOP) networks and ontology frameworks. Data gaps, as identified by AI, will be filled by available state-of-the-art *in vitro* and *in silico* testing methods.

The six NAMs will be evaluated and applied in collaboration with industrial and regulatory stakeholders, in order to maximise end-user acceptance and regulatory confidence. This is anticipated to expedite the implementation of NAMs in risk assessment practices, and facilitate potential regulatory and industrial use.

Descriptive Analysis

Stakeholder Identification

Stakeholders were identified through an iterative process, and were characterised as belonging to one of four groups, namely: i) Government & Policy (i.e. regulatory authorities and policymakers); ii) Industry (i.e. individual companies from various sectors, as well as industry associations); iii) Academia; and iv) Non-governmental Organisations (NGOs) (e.g. animal welfare organisations). During this process, stakeholders were introduced to the details of the ONTOX project and areas of potential collaborative action of mutual interest and benefit to the participating stakeholders; thus, the ONTOX project consortium was formed. Follow-up meetings were organised, to update the respective stakeholders about the progress and challenges of the project.

To determine the expectations and requirements of the individual stakeholders with regard to the project, a stakeholder analysis was performed. This mapping included outlining the strategies of invited stakeholders, their goals and objectives, as well as identifying specific areas of competencies where respective stakeholders would be able to assist ONTOX in achieving its goals. Conversely, ways in which ONTOX could assist the respective stakeholders to achieve their vision, goals and objectives, were also ascertained. The aim of this strategy is to help formulate the emerging ONTOX–stakeholder interaction plan.

The ‘Toxicological Landscape’: Mapping the Current Position of the Stakeholders in Relation to NAMs and PRA

Applied Stepwise Approach. The stakeholder analysis facilitated the identification of relevant stakeholders and

the preparation of targeted presentations. During the one-to-one ONTOX meetings with stakeholders, their views on NAMs, NGRA and probabilistic risk assessment (PRA), were discussed. The results of these discussions resulted in a systematic overview of each stakeholder’s position. A short paper published by ONTOX explains the principles of PRA, including how and when NAMs are meant to be used.²

Survey and Data Analysis. As part of the stakeholder analysis, an online survey was initiated to map the current position of the stakeholders in relation to NAMs and PRA. All 35 identified ONTOX stakeholders (as of 31st January 2023; the number continues to rise) were invited to participate in the survey; 27 stakeholders responded. The purpose of the survey was to identify drivers and barriers towards the implementation of NAMs, NGRA and PRA, and to identify challenges to be discussed in the stakeholder network meeting. The questionnaire consisted of 18 statements where the stakeholders were asked to indicate to which extent they agree or disagree with ratings from 1 (Fully disagree) to 5 (Fully agree). In addition, the stakeholders were asked to elaborate/explain the current position of their organisation in relation to NAMs and PRA, as well as to add anything that was not covered by the 18 statements shown in Table 1. Overall, the 18 statements were clustered into seven topics of concern:

1. Capacity building within and outside your organisation (Statement 9).
2. Sustainability of the platform (requirements) (Statements 10 and 18).
3. Fit of the ONTOX platform in current regulations (Statements 13–17).
4. Validation of AOPs, and acceptance of outcomes (Statements 1, 2 and 4).
5. Quality criteria of data (Statements 6, 7 and 8).
6. Acceptance of AI and machine learning (ML) in risk assessment (Statements 3 and 12).
7. Guaranteeing consumer safety (Statements 5 and 11).

The ratings given to the 18 statements by the stakeholders were used to create polar plots,³ to visualise individual and aggregated stakeholder positions on the seven topics. This data analysis also identified topics on which stakeholders tend to agree and disagree, as can be seen in Figure 1. The data analysis was performed using the R language and environment for statistical computing.⁴ The anonymised data and source of the code can be found in this Github repo: <https://github.com/ontox-project/stakeholder-analysis>.

These survey responses were the basis for organising the detailed agenda for the first stakeholder network meeting, as well as identifying the topics to be addressed in the

Table 1. The list of 18 statements addressed in the ONTOX survey sent to stakeholders in December 2022, regarding their current position in relation to NAMs and PRA.

No.	Statement
1	Mechanistic or biological relevance of adverse outcome pathways (AOPs) is key for the success of the ONTOX approach.
2	All desired toxicological local and systemic effects should be included in the ONTOX approach.
3	Case studies can increase confidence in NAMs and AOPs.
4	It is important to include inter-individual human variability while modelling internal exposure using toxicokinetics in the ONTOX approach.
5	The Threshold of Toxicological Concern (TTC) concept is a way forward.
6	The format in which data are collected and stored must be standardised according to FAIR (Findable, Accessible, Interoperable and Reproducible) principles.
7	The reproducibility crisis has revealed a lack of quality, even in the peer-reviewed literature.
8	In risk assessment, the different parties involved need to speak the same language.
9	The capacity and capability of Contract Research Organisations (CROs) to implement NAMs is insufficient.
10	It is impossible to ensure sustainability if the ONTOX Hub is only designed as an open science platform.
11	Grouping of chemicals is an acceptable approach to reducing the need for testing without hampering confidence in the test results.
12	It is unavoidable to implement machine learning and artificial intelligence in an integrated testing strategy like the ONTOX approach.
13	To accept all results obtained with NAMs without comparing the results with traditional <i>in vivo</i> animal tests, negative results also need to be accepted.
14	The degree of uncertainty that is acceptable for regulation must be defined.
15	The 'one chemical, one risk assessment' ambition can be achieved.
16	Criteria for the use of NAMs in Classification and Labelling of Products (CLP) need to be defined and set up.
17	The ONTOX approach cannot be validated according to the validation processes we have used in the past.
18	The best way to refine the design of the ONTOX Hub is by consulting all stakeholders to obtain their feedback.

The survey's aim was to gauge the current position of stakeholders in relation to non-animal new approach methodologies (NAMs) and probabilistic risk assessment (PRA).

respective breakout groups. The focus here was on topics where the stakeholders had disagreed in their responses, which could represent barriers to the implementation of NAMs and PRA.

Conclusions on the Data Analysis. Overall, analysis of the survey responses identified a number of relevant barriers that must be addressed, as well as some key drivers for the implementation and acceptance of NAMs for chemical risk assessment, NGRA and PRA, according to the goals of the ONTOX project. The analysis also confirmed that there was a variety of opinions within each stakeholder group on the statement that "the 'one chemical, one risk assessment' ambition can be achieved". A need for the firm involvement of all stakeholders from the start of the ONTOX project, so as to ensure a sustainable concept, was also highlighted.

How to Overcome Barriers and Leverage Drivers. The ONTOX approach was designed to facilitate a joint stakeholder network. The ONTOX stakeholder network brings together representatives from scientific communities, regulatory authorities, industry and NGOs, to discuss and develop the concept of PRA. It is envisaged that the stakeholder network

will feature an annual in-person meeting for all relevant stakeholders, with the intention of building a broader end-user acceptance of the NAMs-based PRA of chemicals (end-users in this context are regulatory authorities with responsibility for risk assessment, as well as industry and researchers).

Stakeholder engagement will be arranged in different ways. One follow-up activity is the organisation of a 'hackathon', with the purpose of addressing so-called 'wicked problems'.⁵ In the current context, these problems represent issues raised for which a possible solution was not forthcoming in the first stakeholder network meeting, and where these identified challenges are intended to be tackled by participants outside the consortium. The hackathon will be an event at which a variety of participants will cover a range of different perspectives — i.e. ONTOX members and non-members, academics, regulators, industry, NGOs, and people from other professional areas (e.g. information and communication technologies, finance, creative businesses, energy transition specialisms, and healthcare). The outcomes of the hackathon will represent novel solutions to the identified challenges and issues, the details of which will be presented at the ONTOX stakeholder network meeting, planned for 2024.

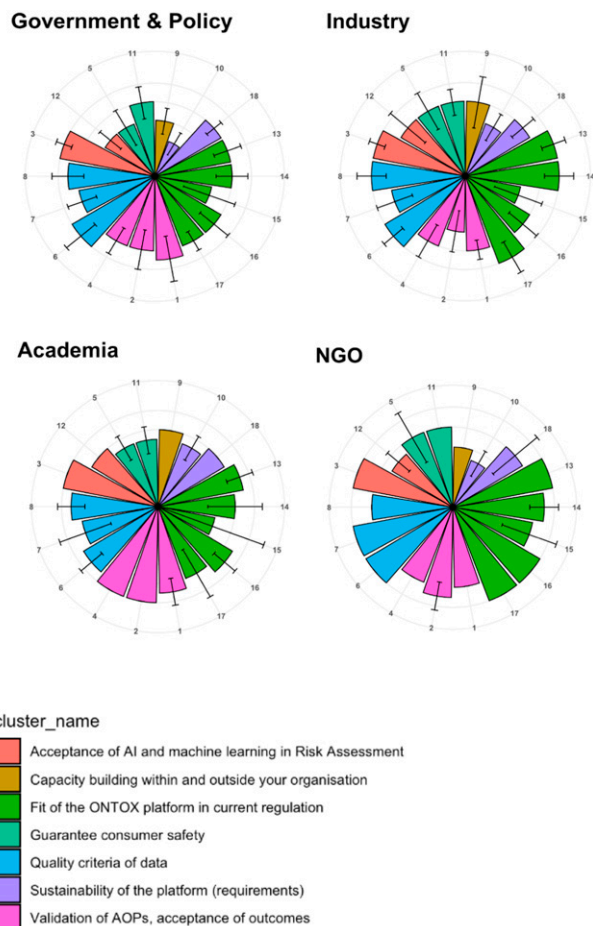


Figure 1. The average scores in the survey, according to the type of organisation. The plots show the average score \pm standard deviation (as error bars), per type of organisation, namely: Government & Policy (10 responders); Industry (13 responders); Academia (2 responders); and NGOs (Non-governmental Organisations) (2 responders). The colours ('cluster_name') indicate the seven different question categories to which the 18 questions in the survey were assigned.

The First Stakeholder Network Meeting

The first ONTOX Stakeholder Network Meeting was held on 13–14 March 2023, in Brussels, Belgium. The overall purpose of the meeting was to map barriers and drivers in relation to the implementation of NAMs and PRA, to address specific issues, and rank these issues from persistent to easy-to-solve. Details on the meeting participants are provided in Table 2.

Participants were organised into six breakout groups, each addressing a key topic regarding NAMs implementation and NGRA. Two facilitators and one rapporteur supported the discussions in each group. Before the meeting, complete instructions for the breakout groups were distributed to all participants. These described the topic content, the planned tasks and roles of the participants/facilitators/rapporteurs, the recommended working schedule (to include a 3-hour discussion time-slot), and the expected output from the breakout group — i.e. a pitch to be presented at a plenary session, followed by a Q&A opportunity and further open discussion.

Output from the breakout groups

In the following sections, reflective summaries of the 3-hour discussions that took place in the six respective breakout groups are presented.

Breakout Group A1: What are the Drivers and Barriers for Full Implementation of NAMs and PRA?

This overall discussion addressed the requirements for the integration of quantitative and qualitative information in the risk assessment process, taking into consideration the role of PRA. The relevant issues discussed were:

Table 2. Participants of the first ONTOX Stakeholder Network Meeting.

Stakeholder group (No. of participants)	Organisations represented
Government and policy (10)	European Commission Joint Research Centre (JRC); Angel Consulting; US Food and Drug Administration (FDA); German Federal Institute for Risk Assessment (BfR); Scientific Committee on Health, Environmental and Emerging Risks (SCHEER); Scientific Committee on Consumer Safety (SCCS)
Industry (12)	Beiersdorf AG; Charles Rivers Laboratories; Cosmetics Europe/ICCS; BASF SE, representing CEFIC; H. Lundbeck A/S; L'Oréal; Novo Nordisk A/S; Symrise AG; Unilever
Academia (2)	Utrecht University/Faculty of Veterinary Medicine/IRAS Tox (The Netherlands); University of Genoa (Italy)
NGOs (2)	Eurogroup for Animals; PETA Science Consortium International e.V.
ONTOX consortium members (24)	The remaining 24 participants were internal ONTOX consortium members, covering all 14 ONTOX work packages.

Invitations were distributed to 38 stakeholders who had indicated their interest in participating in the meeting. Finally, the meeting was held with a total of 50 participants, with 26 participants representing 19 of the 38 invited external stakeholders.

- barriers in relation to the capacity and capability of contract research organisations (CROs) to implement NAMs;
- barriers related to the idea that “PRA is all about uncertainty”;
- defining the criteria for the use of Classification, Labelling and Packaging of substances and mixtures (CLP); and
- the idea that the current validation process should be expanded.

It was stated by the group that PRA is a perfect fit — it integrates different methodologies, and it provides more than ‘Yes or No’ answers. PRA is already applied in different situations (e.g. during an appointment with a doctor, the specialist states the expected efficacy of a proposed therapy). PRA would make risk assessment more transparent through systematic analysis of uncertainty with regard to sensitivity — i.e. some effects are more certain than others. However, there is no 100% certainty in scientific facts. With the implementation of PRA, it would perhaps be possible to abolish currently-used safety factors. Hazard characterisation and exposure were also discussed. The assessment of external versus internal exposure should be a starting point for the risk evaluation process, under the PRA paradigm.

It was concluded that CROs currently do not have sufficient capacity to use NAMs for risk assessment. Adoption of a NAMs-centric PRA paradigm by regulators and policy-makers would create an incentive for CROs to adapt their strategies and portfolios toward this approach. Thus, implementation of NAMs at the regulatory level would be helpful. In addition, data-driven NAMs are complicated, and their use requires considerable expertise and additional training for evaluators. Currently, the value and significance of results from NAMs are weighted differently between various regulatory bodies. Therefore, regulatory harmonisation would be an additional driver for the use of NAMs for risk assessment.

With respect to the validation of NAMs, there are significant differences in how guidance and support are provided from different organisations. Harmonisation of this guidance could provide opportunities for companies and academic groups developing NAMs. Furthermore, the development of guidance documents should be an iterative process, because the field is moving rapidly and thus any documentation should be regularly updated. Case studies are crucial to feed into the harmonisation process, and to demonstrate that integrated testing strategies can work — for example, skin sensitisation was mentioned as a toxicological endpoint where such progress has been made.⁶

Considering CLP, it would help the adoption of NAMs if their positioning in the equivalent CLP-process were to be harmonised between different regions (in this context, regions must be understood in a global perspective).

Discussion Outcome for the Pitch Presentation. This breakout group identified a single challenge — namely, ‘Education and Training’. This was positioned as essential for driving the implementation and acceptance of PRA. The nature of this education would be to deliver training of stakeholders in machine learning, AI, datasets from ONTOX, physiological maps, omics and kinetics data. The outcome of such training should provide insight to participants in the minimum requirements for a dataset/model to give a probability of hazard.

The paradigm shift toward probabilistic reasoning would require a focus on statistical literacy and probability theory, on all levels and for all stakeholders. The expected output of these activities would be differentiated according to the target audience, and would likely take the form of: guidance documents; training materials; public relations materials; practical training; input toward the hackathon conclusions; and scientific research.

In order to obtain a satisfying outcome for this challenge, a number of considerations need to be made. Firstly, the complexity of the challenge needs to be mapped, in order to manage in-scope efforts and focus. The implementation of PRA requires a set of case studies to demonstrate applicability. Data access is important for illustration and reuse purposes, and this should include ONTOX data, a data registry and data package manager (e.g. BioBricks⁷), and QSAR data. An expected timeline for potential strategies was proposed:

- Short-term (end of 2023): Liver injury case study, introduction on PRA.
- Mid-term (April 2026): Identification of a responsible training organisation and the necessary training content.

Overcoming this challenge would be considered to have been successful when there is a relatively large stakeholder acceptance of a specific method, and when it is possible to sufficiently predict a specific toxicological endpoint (such as DILI) with a NAMs-based method implementing PRA.

Breakout Group A2: What are the Drivers and Barriers for Full Implementation of NAMs and PRA?

The overall discussion addressed regulatory acceptance of NAMs — evolution or revolution — and identified a need to be realistic and go with small steps toward testing without the use of animals. Developing a European roadmap, e.g. the European Citizens’ Initiative *Save cruelty free cosmetics — Commit to a Europe without animal testing*,⁸ is important, and discussions with regulators are required. It needs to be recognised that some stakeholders (in industry as well as in academia)

have a high economic interest in continuing the use of animals in experiments. Agencies need time to learn more about the methods, hence they are currently not familiar with NAMs in general. The stakeholder analysis performed by ONTOX has identified significant knowledge gaps, which have also been addressed by the European Commission's initiative to implement training and experience exchange on NAMs for regulators.⁹ There might also be the need for a specific and realistic roadmap for the implementation of NAMs and PRA, and to identify and set specific milestones for this undertaking.

EURL ECVAM belongs to the European Commission's Joint Research Centre (JRC). Its role is not to draft legislation, although it does contribute to legislative proposals by advising the policy Directorate General (DG). Goals and a roadmap are needed (e.g. the International Stakeholder Network; ISTNET¹⁰), in addition to a taskforce for educating end-users of NAMs. Thus, more European Commission investment is needed. Court cases to challenge the regulations, with support from NGOs, might be useful to clarify ambiguities in current regulations (see, for example, *Symrise v. ECHA*¹¹).

It is essential that all stakeholders use the same language to define what we need to achieve, and to help our understanding of the challenges associated with the full implementation of NAMs. The DNT case study (i.e. the Organisation for Economic Co-operation and Development

*Guidance on Evaluation of Data from the Developmental Neurotoxicity (DNT) In-Vitro Testing Battery*¹²) could be used as an example of how to move forward. During the breakout group discussion, specific issues that need to be addressed by the various stakeholder groups were identified, and these are shown in Table 3.

ONTOX partners can arrange talks with regulators, but they cannot solve all of the problems. The NCP (National Contact Point, with reference to *Directive 2010/63/EU*¹³) has access to direct communication with the European Commission and the European Parliament. Each of the ONTOX partners can interact with their NCP and thereby facilitate scientific trust among respective national regulators.

Discussion Outcome for the Pitch Presentation. Two bottlenecks to the implementation of non-animal approaches were identified:

1. The **legislation**, which needs to be changed.
2. The **lack of communication** between industry and regulators.

A number of important characteristics of the identified challenges, as well as some potential solutions, were reported by the group (Table 4). The group recommended that the JRC should take the lead in organising a suggested overall roadmap, and presented a proposal to this end (Figure 2).

Table 3. Specific issues that need to be addressed by the various stakeholder groups.

Stakeholder group	Issues identified
Regulators (e.g. EC, ECHA, EFSA, EMA, EURL ECVAM, national regulators, US EPA, US FDA and OECD)	<ul style="list-style-type: none"> — EC is funding the non-animal approaches but not sufficiently supporting/changing the legislation. — Communication and education are crucial from the start. — There is no overarching organisation for EU agencies. — Laws do not support NAMs and NGRA. — 'Fit-for-purpose' validation is not available. — Targeted development is needed for regulatory purposes. — Prohibitive mandates of agencies are not in place. — There is a lack of confidence and knowledge gaps in the science. — CLP is only hazard-oriented (i.e. not supporting NAMs). — The US EPA and EFSA need to provide funding for the creation of 'fit-for-purpose' NAMs.
Industry	<ul style="list-style-type: none"> — Communication with the EPAA, the European Commission and the European Parliament is crucial. — Improved communication with the regulators is needed. — Funding needs to be put into NAMs (via academia).
NGOs	<ul style="list-style-type: none"> — National-level communication with the regulators is required.

CLP = Classification, Labelling and Packaging of substances and mixtures; EC = European Commission; ECHA = European Chemicals Agency; EFSA = European Food Safety Authority; EMA = European Medicines Agency; EPAA = The European Partnership for Alternative Approaches to Animal Testing; EURL ECVAM = EU Reference Laboratory for alternatives to animal testing/European Centre for the Validation of Alternative Methods; NAMs = new approach methodologies; NGOs = non-governmental organisations; NGRA = next-generation risk assessment; OECD = Organisation for Economic Co-operation and Development; US EPA = Environmental Protection Agency; US FDA = Food and Drug Administration.

Table 4. Characteristics of the identified challenges.

Characteristics of the challenges	Notes relating to potential solutions
<p><i>Challenges can be:</i></p> <ul style="list-style-type: none"> — Highly complex: Multiple stakeholders and policy impact; solvable! — Technical and legislation hurdles. — A revolution within the evolutionary process. — Mechanistic validation: No guidance yet; Case studies are crucial! <p><i>Other aspects include:</i></p> <ul style="list-style-type: none"> — NAMs acceptance is highly complex, making a step-by-step approach necessary for solving this challenge. — CRO involvement depends on regulations. — There is a need to accept a degree of uncertainty. — Complexity toward the community^a is high (involvement of social scientists may provide a solution). 	<p><i>Example strategies:</i></p> <ul style="list-style-type: none"> — Case studies: e.g. DNT roadmap. — Lobbying. <p><i>Expected timeline for a potential solution to be found:</i></p> <ul style="list-style-type: none"> — Depends on the availability of NAMs (e.g. skin sensitisation and DNT could be relatively fast, other fields slower). — There is a need for human resources relative to the expected workload, as well as a coordinator (e.g. JRC) for communication with stakeholders. <p><i>Target audience:</i></p> <ul style="list-style-type: none"> — All invited and participating stakeholders.

CRO = contract research organisation; DNT = developmental neurotoxicity; NAM = new approach methodology; JRC = EU Joint Research Centre.
^aHere, 'the community' refers to, for example: EU agencies, the scientific community, national institutes of public health, national poison centres, on-going research projects dealing with exposure data.

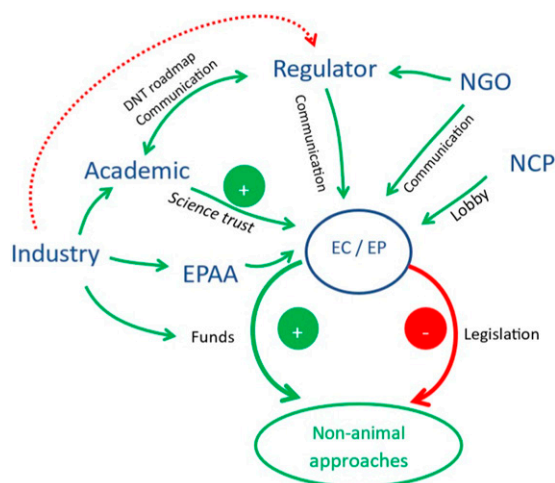


Figure 2. Proposed overall roadmap. EC = European Commission; EP = European Parliament; EPAA = The European Partnership for Alternative Approaches to Animal Testing; NCP = National Contact Point; NGO = non-governmental organisation.

Communication between industry and regulators is crucial. All stakeholders (industry, academia, regulators, NGOs and the NCP) should be invited to participate in the stakeholder network meetings. Also, hackathons that involve fresh target audiences to provide additional new perspectives, can play an important role in finding solutions.

When dealing with regulators, 'evolution' is the best approach, but with small 'revolutions' (e.g. the case study of the DNT roadmap). Those responsible for initiating and driving the roadmap toward NAMs-based risk assessment, including PRA, should be identified as soon as possible.

Changes in the legislation are needed and the EC must be directly addressed. The group would consider the challenges to be successfully overcome when the current legislation has been changed to permit full implementation of NAMs.

Breakout Group B: Hazard Assessment Versus Risk Assessment

There are many unique challenges associated with the use of NAMs. These challenges vary, depending on whether the NAMs are to be used for risk assessment or for hazard assessment, and they also vary according to the industry sector.

The chemical industry sector is more hazard-oriented, and thus hazard-based assessment is high priority. Agrochemical pesticides, active ingredients, formulations and biocides must be included for consideration when implementing NAMs, NGRA and PRA. In contrast, cosmetics regulation is more risk assessment-oriented. For pharmaceuticals, risk and maximum tolerated doses are addressed, but population variability and idiosyncrasy are significant issues.

Quantification of the uncertainty inherent in all methods, including NAMs, is important. Standard operating procedures (SOPs) are needed, in order to ensure reproducibility of non-animal assays, in accordance with universally accepted principles such as 'Good *In Vitro* Method Practices' (GIVIMP).

Proper exposure assessment is challenging for industrial chemicals. In contrast, information about exposure to pharmaceuticals is readily available from pharmacovigilance

reports. Following their discussion, the group collated a list of points to clarify exactly what information is required from such exposure data:

- Susceptibility: This refers to the varying degrees to which different individuals or populations might be affected by exposure to chemicals or pharmaceuticals. Factors like age, genetic predisposition, existing health conditions, and concurrent exposures can influence susceptibility. Understanding susceptibility is crucial in exposure assessment, to ensure that vulnerable groups are adequately protected.
- How safe is ‘safe’? Determining acceptable levels of exposure involves establishing thresholds below which a certain chemical or pharmaceutical is considered safe for most individuals. This determination is critical, as it sets the benchmarks against which actual exposures are measured.
- Human intervention: This involves the role of intentional actions in modifying exposure levels of humans. It could include measures to limit exposure, such as workplace safety protocols, environmental regulations, or prescribing guidelines for pharmaceuticals. These interventions are essential for managing and mitigating exposure.
- Is it hypothesis-driven? This point refers to one of the approaches used in conducting exposure assessments. A hypothesis-driven approach would start with a specific premise or prediction about exposure patterns and then seek data to support or refute this hypothesis. This approach contrasts with data-driven assessments, which start with data collection and then draw conclusions based on observed patterns.
- Occupational exposure: This addresses the exposure of workers to chemicals or pharmaceuticals in the workplace. Occupational exposure assessment is crucial for ensuring workplace safety and health standards, considering workers often face higher and more prolonged exposures than the general population.
- Models for categorisation: This refers to the use of models to predict or categorise exposure levels. Models can be used to estimate exposures in different scenarios, especially where direct measurement is challenging or impractical.
- What are the resources and time requirements? *In vitro/in silico* throughput? This point raises the issue of the resources (e.g. time, money and technical expertise) that are required for different types of exposure assessments.
- Incentives for regulators; how to compensate the risk in generating NAMs: Encouraging regulatory

bodies to adopt innovative methods for exposure assessment is a difficult task, especially in workplaces, where animal studies are seen as the ‘gold standard’. NAMs, however, can offer more efficient, less costly and ethically preferable options, as compared to traditional methods. However, their adoption might require balancing the risks and uncertainties associated with new techniques.

- EU state governance to report activity on alternatives: This point refers to the governance structures within the EU for reporting and monitoring the use of alternative methods in exposure assessment. It emphasises the role of state-level oversight in promoting and ensuring the implementation of safer and more innovative exposure assessment methods.

Discussion Outcome for the Pitch Presentation. How safe is ‘safe’? and Can we live with hazardous substances and protect the planet at the same time? Chemicals are everywhere, some hazardous, others not — indeed, we ourselves are composed of chemicals. Consequently, we need to identify the sources of chemical exposure and the actual dose — i.e. which substances are around us and why (and is it intended or unintended exposure). Criteria with relevance to the identified challenges were identified by the group (Table 5).

The group would consider the challenges successfully solved when ‘Findable, Accessible, Interoperable and Reusable’ (FAIR) databases of exposure data, *in silico* models for various exposure scenarios, and Environmental Protection Agency (EPA) database-like tools (plug-in for PRA) are available. However, the time frame for successfully overcoming the challenge will be sector-dependent — i.e. it will likely be sooner in sectors that have already adopted some risk assessment approaches instead of relying solely on hazard assessment.

Developing and using relevant exposure scenarios could mitigate the limited availability of exposure data. For most industrial chemicals, robust and reliable exposure data are very limited and benchmarking with, for example, pharmaceuticals and biocides could be beneficial. Uncertainty of exposure data must be addressed through a focus of attention on finding a solution to the lack of exposure data for industrial chemicals. Connecting with the Safe and Sustainable by Design (SSbD) concept,¹⁴ as well as the OSOA (i.e. one substance, one assessment) concept,¹⁵ could bridge this gap, as exposure (occupational and consumer) is central to both.

Breakout Group C: How to Promote Future Reliance on Artificial Intelligence (AI) and Machine Learning (ML)

This breakout group addressed the narratives generated by the stakeholders when elaborating upon their current

Table 5. Criteria for the selected challenges.

Criterion	Notes
Challenge complexity and degree of solvability	<p><i>Dependent on:</i></p> <ul style="list-style-type: none"> — Availability of exposure data; — Uncertainty, i.e. the lack of complete knowledge or precision in the available data and methodologies used to assess and characterise the risks associated with exposure to certain substances; — Resources/time; — Susceptibility, i.e. the varying degrees to which different individuals or populations might be affected by exposure to chemicals or pharmaceuticals. Factors such as age, genetic predisposition, existing health conditions, and concurrent exposures can all influence susceptibility. Understanding susceptibility is crucial in exposure assessment, to ensure that vulnerable groups are adequately protected; — Regulatory context.
Data access	<ul style="list-style-type: none"> — EU member state dependent; — Can be subject to GDPR issues, language differences and FAIR issues, as well as issues over fees and data quality.
Expected timeline for a potential solution to be found	<ul style="list-style-type: none"> — Sector-dependent, i.e. depends on the availability of reliable exposure data; — Need for human resources, relative to predicted workload — e.g. long-term EU-sponsored projects such as ‘Horizon Europe’.
Target audience	<ul style="list-style-type: none"> — The invited and participating stakeholders.
Complexity toward the community ^a	<ul style="list-style-type: none"> — Complexity level of the challenge in this respect should be medium/high.

FAIR = Findable, Accessible, Interoperable and Reusable; GDPR = General Data Protection Regulation.

^aHere, ‘the community’ refers to, for example: EU agencies, the scientific community, national institutes of public health, national poison centres, on-going research projects dealing with exposure data.

positions in relation to NAMs and PRA (see Table 1), and sought to identify the associated challenges relevant to AI and ML. The identified challenges were:

- addressing the ‘explainability’ of AI;
- defining the applicability domain of AI in toxicological risk assessment;
- adapting the process of risk assessment to include AI;
- adapting PRA to include AI; and
- building trust in AI systems.

Addressing the ‘Explainability’ of AI. This term has been coined in the field of computational sciences. It can be very difficult, particularly in the case of highly recursive and very deep learning AI models, to explain how AI predictions are made. Several issues that need to be addressed concerning this topic were identified. For example, in the context of toxicological risk assessment, the user’s level of understanding of AI, and the information (data) requirements, will need to be considered, i.e.:

- Information (data): hyperparameter settings; data used to train the model; model type and

settings; data flowing in and out of the AI models need to be traceable and described in documentation.

- Understanding of AI: a solid understanding of how AI works will be required by the people relying on it; training and education is needed for increasing statistical literacy and the general understanding of AI models.

Defining the Applicability Domain of AI. The question: *Are we intending to use AI to provide us with information, or will we rely on AI to make decisions?* is relevant in this context.

In general, the breakout group was convinced that the broad applicability of AI will be as an information-delivery tool. A further question was also considered by the group, namely: *Will AI tools for risk assessment be expert systems, or will they be available to a more general user-base?* The meaning of ‘expert system’ in this context refers to the users being ‘experts’ in AI. Thus, users would need to know about AI and how it operates, in order to work with it. The system could also be considered an ‘expert system’ with respect to toxicology, as it would be necessary to understand the limits of AI and also be able to understand which toxicological questions can be answered. However, the group discussion

focused on building systems that are, in essence, ‘expert systems’ that are not so strictly constrained, and therefore they would be available to the more general user. Thus, with proper instruction and with some domain knowledge, the general user would be able to interact with the systems effectively and in a safe manner. In this respect, education and training were mentioned as crucial in helping the general user to interact with the systems properly. The ability to experiment with the systems at an early stage of learning, and also to gain insight into how certain predictions are produced by AI systems, were also considered important points (i.e. the ‘explainability’ of AI systems).

Adapting the Process of Risk Assessment to Include AI. Considering predictions, AI systems may generate ‘red flags’ for compounds/compound groups not previously flagged up by other approaches. How will authorities, industry and society deal with this? The current risk evaluation process seems self-supporting: requirements for risk assessment are anchored in the law, by means of strict guidelines, and industry is bound by this legislation. How will this transformation of the risk assessment paradigm take place, whilst ensuring protection of the population, individuals or at-risk groups? Will the AI-driven methods help, or will they only complicate the paradigm? No clear answers were formulated, but a way forward would be to increase AI ‘explainability’ and the traceability of the input/output. Furthermore, in all cases, the responsibility for the decision-making should reside at the human level. Reshaping the validation framework to include AI-driven methods, would be a good first step.

Adapting PRA to Include AI. Under the current risk assessment framework, there is little room for a probabilistic approach. The determined risk is a black, grey or white answer. The integration of AI into the risk assessment paradigm can be a chance to update our beliefs — meaning that instead of arriving at point estimates for risk assessment metrics, we can get probability distributions for those metrics (or some proxy measure for that), based on population characteristics or scenarios. In some cases, the centre measure from this distribution (e.g. the mean) will be sampled from a very broad distribution. This means, in fact, that the certainty of that centre measure is low. In the case where there is a very narrow distribution, the centre measure sampled from that distribution has quite a high certainty. This means that the answers will not be grey, but will be surrounded by estimates of (un)certainty of that measure being representative of the true measure. It will yield a way to describe how sure/unsure we are about a specific measure, and probably can lead to new, more focused research, to address the uncertainty and strive to make it smaller. AI models are probabilistic machines and inherently AI provides uncertainty about given predictions.

Building Trust in AI Systems. For AI systems to gain trust, it would be necessary to create, share and demonstrate case studies of their successful use, leading to increased experience in using AI for risk evaluation purposes — ONTOX could deliver on this. Additional barriers that could limit or inhibit the implementation of AI-driven methods in toxicological risk assessment, and some suggestions for their mitigation, were identified:

- The true value of AI-driven methods in toxicological risk assessment should be demonstrated in bridging studies, using existing compounds/compound groups (for example, making a stronger case for the mechanistic understanding of the impact of perfluorinated and polyfluorinated alkyl substances (PFAS) on human health), when we stack AI-derived predictions on what we already know in this case.
- What would really drive AI adoption? Would this be a good topic for a hackathon?
- How would we deal with transforming the current paradigm where decision making relies on implicit weight-of-evidence (WoE)? This is where AI could be superior in terms of objectiveness, but it is not clear how to evaluate/validate AI methods against something that is not transparent.

So how can trust in AI methods be increased? Currently, it is a technical challenge to ‘open the box’ on AI. In theory, explaining the principles of the inner workings of deep learning models is fairly straightforward. To show how sophisticated deep learning AI models precisely arrive at their predictions can be a lot harder. It is a heavily pursued scientific field that has gained a lot of traction over the past five years. With the arrival of large language models, such as OpenAI’s GPT-3.5 and its successor GPT-4,¹⁶ the call for increased explainability has become stronger, from societal, business and scientific perspectives. A solution could be to leverage AI models’ intrinsic potential for reasoning on their applicability domain and limitations. This could work particularly well for generative large language models. The current models already display some notion on how well they are able to perform a task (for example, refer to the blog: <https://medium.com/@enolve.io/i-had-chatgpt-tell-me-about-its-inner-workings-heres-what-i-learned-d5c0f211b7b5>). A generic solution for deep learning models was proposed by Lew et al.¹⁷ at the 26th International Conference on Artificial Intelligence and Statistics (AISTATS), which was held in Valencia, Spain, in April 2023. Here, they showed that, by using a specific type of Monte Carlo simulation, it was possible to extract from a deep learning model information on how good it estimates its predictions to be.

Discussion Outcome for the Pitch Presentation. Transparency and lack of insight hampers trust and acceptance. The breakout group identified that ‘pulling the lid off’ AI may be necessary, to improve transparency of the process of risk assessment by moving from argumentation to explanation (mechanism-driven), and to help address unfamiliarity. Possible solutions include: providing education/training; working on demo cases; and defining quality criteria — for example, standards and SOPs.

As a challenge for the hackathon, the group proposed formulating a task to identify the drivers of acceptance of NAMs, AI and machine learning for regulatory purposes, as well as to determine the factors hampering the acceptance of these methods. Criteria by which to define outcome and acceptance are yet to be determined.

Breakout Group D: How to Make Future Chemical Testing Sufficient, and the Requirements for Industry Achievable?

The discussion addressed challenges in the industrial uptake of NAMs. Firstly, it is difficult to implement NAMs in the absence of incentives or clear regulatory requirements for the industry. Secondly, the transition to NGRA is hampered by the lack of reliable NAM-based approaches for the more complex toxicological endpoints, such as repeated dose and reproductive toxicity. At the same time, the stakeholders generally agreed that NAMs are already in place for local toxicity and are used, to different extents, in regulatory hazard and risk assessment.

Regulators argued for incentives to conduct NAM-based and traditional *in vivo* testing in parallel for a certain period, during a so-called ‘transition phase’. Such a transition phase could generate the required experience (and confidence) in the NAM-based safety assessment of repeated dose toxicity or reproductive toxicity. Simultaneously, there is a growing demand for new products — and these can contain chemicals that are challenging to test *in vitro*, and which fall outside the applicability domain of *in silico* NAMs, e.g. peptides or nanomaterials. Furthermore, the recent development of the EU CLP regulation, which requires an assessment of the endocrine-disrupting properties of chemicals, poses new challenges to industry and risk assessors alike.

Overall, the breakout group agreed that there is a need to develop systemic toxicity NAMs, as targeted by the ONTOX project.

Discussion Outcome for the Pitch Presentation. The ONTOX project partially addresses the need for NAMs applicable to complex endpoints, such as those envisioned by work packages dealing with liver toxicity, nephrotoxicity and effects on neuronal tube closure.

The stakeholders classified each of the challenges associated with the industrial uptake of NAMs under one of the following key issues:

1. *Validation:* Scientifically valid NAMs; interlaboratory transferability.
2. *Translation:* Uptake from academia and transfer of NAMs into the GLP/CRO environment.
3. *Regulation:* NAMs are fully accepted and promoted, regulations are aligned.
4. *Education:* Stakeholders learn how to guarantee consumer safety by using NAMs.

However, one should not rely on the replacement (of established *in vivo* methods) alone. The group proposed to stimulate the development of Integrated Approaches to Testing and Assessment (IATA) approaches, with the goal of minimal animal-use in combination with NAMs to allow for a smooth transition.

As a challenge for the hackathon, the group proposed the development of a plan for data gap filling for systemic toxicity and/or for specific organ toxicities, to include the development, implementation and acceptance of NAMs. Thus, the plan might include the use of:

- NAMs or NAM combinations (e.g. IATAs) that are applicable to ‘real world’ scenarios, for the testing of, for example, environmental chemicals (not bioactive by design) and difficult chemicals (peptides, nanomaterials, polymers, EDCs, etc.), **leading to**
- industrial uptake of NAMs; regulatory acceptance of NAMs; and strengthening of the NAMs culture in the stakeholders’ environments.

The group would consider the challenges to be successfully overcome when the regulatory acceptance of newly-developed IATAs with NAMs is achieved, and the use of animal-free NAMs is implemented successfully and smoothly. In the transition phase, available studies on chemicals with animal data (e.g. OECD Test Guidelines) should be benchmarked with studies based on the ONTOX NAMs and PRA approach, to demonstrate that (hopefully) the same conclusions are reached. As for the pharma and cosmetic sector, one should have learnt how to guarantee consumer safety by using the new methods.

Breakout Group E: Sustainability of ONTOX Hub: End-User Challenges

The overall discussion in this breakout group addressed how to ensure that ONTOX Hub will add value in the future and will become the toxicological tool of choice. ONTOX Hub aims to be a data repository, with information to be used by risk regulators, and to provide direct

probabilistic risk assessment output. The ONTOX Hub platform will thus serve three key roles:

- a knowledge framework to provide end-users with access to the *in silico* tools developed in the ONTOX project;
- a marketplace for the NAMs (including *in vitro* assays and *in silico* tools) developed in the ONTOX project; and
- a resource for accessing the specific expertise available in the consortium, including training and consulting services.

The intention is for ONTOX Hub to be an intuitive one-stop shop, avoiding the need for multiple tools and thus saving time and resources. It will include access to additional information for risk assessment, while providing different levels of information for the various end-users. For example, regulators could have access to all the data, while other users would have access to only some of the data. Options on the dashboard could accommodate various levels of expertise — from layperson to expert level — and signposts could be provided to facilitate digging deeper into the data, if needed. End-users would have the option of either performing a deterministic risk assessment, or

collecting information for a fully probabilistic risk assessment. Direct access to additional models and databases from outside the ONTOX project could be added within ONTOX Hub, to strengthen probabilistic risk assessment. The dashboard must be intuitive and include a manual, a help function and a feedback function. Organised training for users must be offered.

To ensure confidence in the results, ONTOX Hub must not be a ‘black box’ and therefore transparency, accessibility and education are key. Transparency about input data (currently, data are coming from the COSMOS database, but more will be incorporated), the applied algorithm and output data must be provided. All the datasets need to be available, at least to regulators, as well as the applicability domain (where ONTOX Hub clarifies the chemical and biological space). Specific examples (case studies) can help to build confidence.

For ONTOX Hub to become the future toxicological tool of choice, its long-term sustainability needs to be completely ensured. The importance of hosting ONTOX Hub post-project has already been established — including the need for user training and support. During the group discussion, it was put forward that if it is hosted and maintained by an agency, it might be more easily accepted (e.g. the EFSA toolbox) than if it is maintained as a ‘paid-for’ system.

Table 6. Recommended issues for the proposed hackathon.

Breakout group	Topic addressed	Recommended issues for the hackathon
A1	What are the drivers and barriers for full implementation of NAMs and PRA?	Education and training on PRA of end-users. — Defining content and responsibility.
A2	What are the drivers and barriers for full implementation of NAMs and PRA?	Development of a roadmap for implementation of NAMs and PRA. — Defining responsibility for taking leadership, involvement of stakeholders, content and ambition.
B	Hazard assessment versus risk assessment. <i>How safe is ‘safe’?</i> <i>Can we live with hazardous substances and protect the planet at the same time?</i>	The development of relevant exposure scenarios and make them available. — Addressing uncertainties related to exposure data and scenarios.
C	How to promote future reliance on Artificial Intelligence (AI) and Machine learning (ML).	A focus on drivers of acceptance of AI, ML for regulatory purposes, and what is hampering acceptance. — Improving transparency, ‘explainability’ and trust of AI within the context of risk assessment.
D	How to make future chemical testing sufficient and the requirements for industry achievable.	Systemic toxicity and data gap filling. — Identifying drivers of acceptance of NAMs, AI and ML for regulatory purposes, and what is hampering acceptance.
E	Sustainability of ONTOX Hub: End-user challenges.	What is the added value of the ONTOX Hub platform? — Defining the requirements for a sustainable platform.

AI = artificial intelligence; NAMs = new approach methodologies; ML = machine learning; PRA = probabilistic risk assessment.

Backing from the regulatory agencies will also be important to promote its use. In addition to maintenance, ONTOX Hub requires a strategy for model improvement by continuously integrating new data and external model updates. In April 2026, the ONTOX cluster projects will be concluded — therefore, the minimum requirements for ONTOX Hub to be considered ‘ready’ must be defined as soon as possible.

Discussion Outcome for the Pitch Presentation. Building confidence in the ONTOX Hub platform requires transparency and visibility. This can be achieved by defining what it can and cannot do (e.g. PRA), as well as the applicability domain/chemical and biological space. The level of detail of the information made available to the different users, starting with the bare minimum, must be defined. For example:

- the training dataset;
- user training and support;
- validation (internal and external); and
- case studies.

Compared with other tools, ONTOX Hub might represent added value for the project and beyond, if it offers the following:

1. A ‘Bucket of tools’.
2. A ‘One-stop shop’.
3. Expert support and guidance for users.
4. Long-term sustainability.

Conclusion and next steps

The first ONTOX Stakeholder Network Meeting identified several issues which must be addressed on the way toward a full implementation of NAMs and PRA. One of the meeting outcomes was a list of recommendations for issues to be addressed during a hackathon in 2024. These are summarised in [Table 6](#).

ONTOX will welcome potential solutions to these various complex issues during the hackathon, planned to take place in April 2024. The hackathon will involve participants from industry, government, academic organisations and NGOs. Additionally, a unique feature of the hackathon will be the involvement of participants from a range of diverse backgrounds, who will be able to offer different perspectives. This will enable a creative ‘out-of-the-box’ approach to overcoming the challenges identified by ONTOX, benefitting considerably from the available interdisciplinary expertise.

In the interim, ONTOX aims to identify the most suitable proposals for the hackathon challenges and finalise the details of the event.

Acknowledgements

The ONTOX project received funding from the European Union’s Horizon 2020 Research and Innovation programme, under Grant Agreement No. 963845.

Disclaimer

The opinions expressed in this publication do not necessarily represent the regulatory position of the institutes or organisations whose representatives participated in the expert meetings and discussions.

ORCID iDs

Michael Santillo  <https://orcid.org/0000-0001-8087-5859>
 Denise Bloch  <https://orcid.org/0000-0002-1222-9668>
 Olena Kucheryavenko  <https://orcid.org/0000-0002-6349-4824>
 Abdulkarim Najjar  <https://orcid.org/0000-0003-4620-9830>
 Predrag Kucic  <https://orcid.org/0000-0002-0597-8927>
 Merel Ritskes-Hoitinga  <https://orcid.org/0000-0001-5315-284X>

References

1. ONTOX. *Homepage*, <https://ontox-project.eu/> (2023, accessed 24 November 2023).
2. ONTOX. *Introduction to probabilistic risk assessment by ONTOX*, <https://ontox-project.eu/introduction-probabilistic-risk-assessment/> (2023, accessed 24 November 2023).
3. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag, 2016, 213 pp.
4. The R Foundation. *The R Project for Statistical Computing*, <https://www.r-project.org/> (2023, accessed 24 November 2023).
5. Stony Brook University. *What’s a Wicked Problem?* <https://www.stonybrook.edu/commcms/wicked-problem/about/What-is-a-wicked-problem> (2023, accessed 24 November 2023).
6. European Commission. *TSAR — Tracking System for alternative methods towards regulatory acceptance: Genomic Allergen Rapid Detection test*, <https://tsar.jrc.ec.europa.eu/test-method/tm2011-09> (2023, accessed 24 November 2023).
7. BioBricks Foundation. *Building with biology to benefit all people and the planet*, <https://biobricks.org/> (2023, accessed 24 November 2023).
8. European Citizens’ Initiative. *Save cruelty free cosmetics — Commit to a Europe without animal testing*, https://citizens-initiative.europa.eu/initiatives/details/2021/000006_en (2023, accessed 24 November 2023).
9. European Commission. *Funding & tender opportunities: Find calls for proposals and tenders*, <https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/home> (2023, accessed 24 November 2023).
10. Bal-Price A, Crofton KM, Leist M, et al. International STakeholder NETwork (ISTNET): Creating a developmental neurotoxicity (DNT) testing road map for regulatory purposes. *Arch Toxicol* 2015; 89: 269–287.
11. Curia. *InfoCuria: Case-law. List of results: T-656/20 — Symrise v. ECHA*, <https://curia.europa.eu/juris/liste.jsf?lgrec=>

- [fr&td=%3BALL&language=en&num=T-656/20&jur=T](#) (2023, accessed 24 November 2023).
12. OECD. *Guidance on Evaluation of Data from the Developmental Neurotoxicity (DNT) In-Vitro Testing Battery Original Draft: July 2021 Revised Draft: October 2022*. Paris, France: Organisation for Economic Co-operation and Development, 2022, 48 pp.
 13. European Parliament. *Directive 2010/63/EU* of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes. *Off J Eur Union* 2010; 20: 33–79.
 14. Patinha Caldeira C, Farcal R, Moretti C, et al. JRC Technical Report. Safe and Sustainable by Design chemicals and materials. *Review of safety and sustainability dimensions, aspects, methods, indicators, and tools, EUR 30991 EN*. Luxembourg: Publications Office of the European Union, 2022, 184 pp.
 15. EFSA. *In support of the EU chemicals strategy for sustainability: One substance — one assessment*, <https://www.efsa.europa.eu/en/corporate/pub/oso> (2020, accessed 24 November 2023).
 16. OpenAI. *Models*, <https://platform.openai.com/docs/models> (2023, accessed 24 November 2023).
 17. Lew AK, Matheos G, Zhi-Xuan T, et al. SMCP³: Sequential Monte Carlo with probabilistic program proposals. *PMLR* 2023; 206: 7061–7088.

Appendix

Table AI. Abbreviations used in the main text.

AI	Artificial Intelligence
CLP	Classification, Labelling and Packaging of substances and mixtures
CRO	Contract research organisation
DILI	Drug-induced liver injury
DNT	Developmental neurotoxicity
EC	European Commission
EC JRC	EC Joint Research Centre
ECHA	European Chemicals Agency
EFSA	European Food Safety Authority
EP	European Parliament
EPAA	The European Partnership for Alternative Approaches to Animal Testing
EMA	European Medicines Agency
EURL ECVAM	EU Reference Laboratory for alternatives to animal testing/European Centre for the Validation of Alternative Methods
FAIR	Findable, Accessible, Interoperable and Reusable
GDPR	General Data Protection Regulation
GIVIMP	Good <i>In Vitro</i> Method Practices
GLP	Good Laboratory Practice
IATAs	Integrated Approaches to Testing and Assessment
NAMs	New Approach Methodologies
NCP	National Contact Point
NGO	Non-Governmental Organisation
NGRA	Next Generation Risk Assessment
OECD	Organisation for Economic Co-operation and Development
PRA	Probabilistic Risk Assessment
QSAR	Quantitative Structure–Activity Relationship
RA	Risk Assessment
REACH	Regulation on the Registration, Evaluation, Authorisation and Restriction of Chemicals
SCCS	Scientific Committee on Consumer Safety
SCHEER	Scientific Committee on Health, Environmental and Emerging Risks
SOP	Standard Operating Procedure
SSbD	Safe and Sustainable by Design
TG	Test Guideline
US EPA	US Environmental Protection Agency
US FDA	US Food and Drug Administration
WoE	Weight-of-Evidence