Downloaded from orbit.dtu.dk on: May 10, 2024

DTU Library

Advancing exposure data analytics and repositories as part of the European Exposure Science Strategy 2020−2030

Kosnik, Marissa B.; Kephalopoulos, Stylianos; Muñoz, Amalia; Aurisano, Nicolò; Cusinato, Alberto; Dimitroulopoulou, Sani; Slobodnik, Jaroslav; De Mello, Jonathas; Zare Jeddi, Maryam; Cascio, Claudia

Total number of authors:
14

Published in:
Environment International

Link to article, DOI:
10.1016/j.envint.2022.107610

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Kosnik, M. B., Kephalopoulos, S., Muñoz, A., Aurisano, N., Cusinato, A., Dimitroulopoulou, S., Slobodnik, J., De Mello, J., Zare Jeddi, M., Cascio, C., Ahrens, A., Bruinen de Bruin, Y., Lieck, L., & Fantke, P. (2022). Advancing exposure data analytics and repositories as part of the European Exposure Science Strategy 2020−2030. *Environment International*, *170*, Article 107610. https://doi.org/10.1016/j.envint.2022.107610

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Environment International 170 (2022) 107610

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

Advancing exposure data analytics and repositories as part of the European Exposure Science Strategy 2020–2030

Marissa B. Kosnik [a], Stylianos Kephalopoulos [b,*], Amalia Muñoz [c], Nicolò Aurisano [a], Alberto Cusinato [b], Sani Dimitroulopoulou [d], Jaroslav Slobodnik [e], Jonathas De Mello [f], Maryam Zare Jeddi [g], Claudia Cascio [h], Andreas Ahrens [i], Yuri Bruinen de Bruin [j], Lothar Lieck [j], Peter Fantke [a,*]

[a] Quantitative Sustainability Assessment, Department of Environmental and Resource Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
[b] European Commission, Joint Research Centre (JRC), Ispra, Italy
[c] European Commission, Joint Research Centre (JRC), Geel, Belgium
[d] Air Quality and Public Health, EHE Dept, UK Health Security Agency, Chilton OX11 0RQ, United Kingdom
[e] Environmental Institute, Okruzna, 784/42, 972 41 Kos, Slovak Republic
[f] Economy Division, United Nations Environment Programme, 75015 Paris, France
[g] National Institute for Public Health and the Environment (RIVM), 3721 MA Bilthoven, the Netherlands
[h] European Food Safety Authority (EFSA), Parma, Italy
[i] European Chemicals Agency (ECHA), Helsinki, Finland
[j] European Agency for Safety and Health at Work (EU-OSHA), Bilbao, Spain

ARTICLE INFO

Handling Editor: Adrian Covaci

Keywords:
International Society of Exposure Science
Data analysis
Data management
Chemicals
Exposure assessment
Risk assessment

ABSTRACT

High-quality and comprehensive exposure-related data are critical for different decision contexts, including environmental and human health monitoring, and chemicals risk assessment and management. However, exposure-related data are currently scattered, frequently of unclear quality and structure, not readily accessible, and stored in various—partly overlapping—data repositories, leading to inefficient and ineffective data usage in Europe and globally. We propose strategic guidance for an integrated European exposure data production and management framework for use in science and policy, building on current and future data analysis and digitalization trends. We map the existing exposure data landscape to requirements for data analytics and repositories across European policies and regulations. We further identify needs and ways forward for improving data generation, sharing, and usage, and translate identified needs into an operational action plan for European and global advancement of exposure data for policies and regulations. Identified key areas of action are to develop consistent exposure data standards and terminology for data production and reporting, increase data transparency and availability, enhance data storage and related infrastructure, boost automation in data management, increase data integration, and advance tools for innovative data analysis. Improving and streamlining exposure data generation and uptake into science and policy is crucial for the European Chemicals Strategy for Sustainability and European Digital Strategy, in line with EU Data policies on data management and interoperability.

1. Introduction

The volume of data supporting chemical monitoring and human/environmental exposure and risk assessment processes linked to various European Union (EU) regulatory, science, and policy making contexts is growing at an unprecedented pace, with a continuously expanding body of scientific literature, and numerous growing data repositories (databases and data platforms) (European Commission, 2020a). While more data and scientific progress are clearly beneficial, it is increasingly difficult to find, compare, combine, and reuse relevant information, keep up with research developments globally, and make use of the fast growing body of associated data (Chu and Evans, 2021). Humans are exposed to an ever-growing number of marketed chemicals in a variety of complex exposure settings (Carney Almroth et al., 2022; Jolliet et al.,

* Corresponding authors.
E-mail addresses: stylianos.kephalopoulos@ec.europa.eu (S. Kephalopoulos), pefan@dtu.dk (P. Fantke).

https://doi.org/10.1016/j.envint.2022.107610
Received 30 May 2022; Received in revised form 28 October 2022; Accepted 29 October 2022
Available online 30 October 2022
0160-4120/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

2021). Global society faces various challenges associated with environmental stressors, from chemical pollution to pathogens such as SARS-CoV-2 (Bruinen de Bruin et al., 2020), and related pressure on human health and environmental quality (Groh et al., 2022; Johnson et al., 2020; Landrigan et al., 2018; Liu et al., 2021). All these are drivers for a growing need to harness existing and emerging data, information, and technologies for sustainable management of chemicals and other stressors. With an aim to protect human health and environmental quality in Europe and at a global scale (Backhaus et al., 2018; United Nations Environment Programme, 2019a), improved exposure-related data generation and management would support ongoing national and international strategies. Examples include European action plans for a non-toxic environment (European Commission, 2017), Europe's Beating Cancer Plan (European Commission, 2021a), the Strategic Approach to International Chemicals Management (SAICM, 2020), and a viable circular economy (Euopean Commission, 2015; United Nations Environment Programme, 2019b).

We have a unique opportunity for advancing exposure-related data analytics and repositories in Europe. This opportunity is driven by the objectives of the European Chemicals Strategy for Sustainability (European Commission, 2020b) (as part of the European Green Deal (European Commission, 2019a)), which strives—among other things—for increased accessibility and digitalization of data production and flows related to chemicals along their entire life cycles, and developing better cross-sector cooperation and applicable assessments. It is also driven by the European Digital Strategy, which includes a concept for promoting Artificial Intelligence (AI) for Europe (European Commission, 2018a) and the European Data Strategy (European Commission, 2020a), aiming to improve Europe's competitiveness around data generation and management, including exposure-related data, by promoting funding for new data infrastructures with long-term sustainability strategies.

Exposure assessment is a critical component for sound chemicals risk assessment and management, with different types of exposure data needed across various regulatory and non-regulatory frameworks to evaluate related risks and impacts, public and environmental health, as well as food, energy, and environmental safety and sustainability (Fantke et al., 2020b). Exposure-related data comprise monitoring data for chemicals and other stressors in various media/environments (e.g., humans, ecosystems, food/feed, consumer products, workplaces), modelled external and internal exposure concentrations, stressor occurrence and use data (e.g., consumer behaviour, food consumption, product use), data on chemical metabolism in organisms, exposure-related physicochemical properties, and chemical production volumes and emissions as exposure sources. These data are sought by various actors and stakeholders (policy makers, industry, academia, regulators, and the general public) for use across EU policies and regulations with synergistic developments and data cross-fertilisation done in collaboration with international organizations/bodies (e.g., World Health Organization (WHO), United Nations (UN), the Organization for Economic Cooperation and Development (OECD)). Several major exposure-related data generation initiatives have been established in Europe such as the HBM4EU project for producing and collecting human biomonitoring data (Ganzleben et al., 2017), as well as parallel exposure data storage and management efforts. A variety of data types are covered by repositories, such as the database of registered substances information submitted under the European REACH (Registration, Evaluation, Authorisation and Restriction of Chemicals) Regulation (European Commission, 2006), the European Chemical Agency's (ECHA) SCIP (Substances of Concern In Products) database established under the European Waste Framework Directive (European Commission, 2008), the Comprehensive Food Consumption database (EFSA, 2011) and Chemical Hazards Database (OpenFoodTox) (Dorne et al., 2021) of the European Food Safety Authority (EFSA), the Information Platform for Chemical Monitoring (IPCHEM) operated by the European Commission in collaboration with European Union Agencies (Joint Research Center, 2020; Kephalopoulos et al., 2020) and the European database on occupational exposure to hazardous chemicals (HazChem@Work) (European Commission, 2016). However, Europe faces specific challenges and conditions for exposure-related data generation, management, sharing, and use/reuse that are related to a complex regulatory framework, including scattered databases, data that are often unpublished or not readily available, and a lack of a common glossary owing to sectorial approaches (Bruinen de Bruin et al., 2022; Fantke et al., 2020b).

Despite a global trend to boost data-driven and digitalization approaches in industry and regulation (European Commission, 2021b; Fantke et al., 2021), reliable and high-quality exposure data are still only available for a small fraction of marketed chemicals and limited exposure settings (Kristiansson et al., 2021), thus hampering advances in exposure science and related uptake in EU chemical policy, regulations, and other relevant decision support contexts (Bruinen de Bruin et al., 2022). Furthermore, even when data exist, they are often not fully available or are restricted to specific groups of end users. For example, only a fraction of the European Chemical Agency's (ECHA) modelled exposure data are publicly disseminated, limiting their use beyond REACH (Fantke et al., 2020a; Müller et al., 2017).

There is a need to link efforts in exposure-related data generation and integration with an overarching exposure science strategy aligned with European and global sustainable development targets for managing chemical pollution. As advocated by one of the five key areas of the 'Europe Regional Chapter of the International Society of Exposure Science' (ISES Europe) Working Group on 'Integrated Framework of Exposure Science and Policy Efficiency' this can be achieved by creating a shared scientific framework for exposure assessment that interfaces with EU chemical policies for health, environment, and risk and sustainability assessment (Bruinen de Bruin et al., 2022). To support and advance exposure science and its uptake into European and global policies and regulations, this translates into a need for a comprehensive high-quality data collection, curation, and integration framework, supported by a harmonised process of generating, collecting, reporting, and sharing exposure-related data sources, methods, and tools for data analysis. Following up on the foundations of the European Strategy on Exposure Science 2020–2030 established by ISES Europe in 2019 (Fantke et al., 2020b), the aim of the present study is to propose strategic guidance for such an integrated framework of European exposure data production, management, and usage. While the focus of our work considers the unique European challenges around regulatory structures and data collection priorities as part of the European Strategy efforts, we also aim to foster alignment with global efforts.

In order to achieve our goal, the ISES Europe 'Working Group on Exposure data repositories and analytics' focused on three specific objectives:

(a) To provide an overview and map existing exposure data onto related requirements for data analytics and repositories across European regulations,

(b) To identify needs and ways forward for improved exposure data generation, management, and usage, and.

(c) To translate the identified needs into operational actions, increase regulatory uptake of (evidence-based) exposure data in Europe, and outline an approach for advancing, integrating, and aligning European data collection, management, and multi-use efforts with global initiatives.

The resulting action plan for exposure data analytics and repositories is one of the priority areas of the 'European Exposure Science Strategy' 2020–2030 (Fantke et al., 2020b), aligns with the vision of "Europe's Digital Decade" (European Commission, 2021b), and constitutes an important step toward global harmonisation of exposure data usage in support of the European Commission's 'One substance, one assessment' approach as part of the European Chemical Strategy for Sustainability (European Commission, 2020b).

2. The European exposure data landscape

Tailored exposure data and tools need to be developed that meet the varied needs of different EU policies and regulations while facilitating cross-utilisation across the policy landscape. Numerous challenges can hinder exposure data uptake and use by practitioners, including finding relevant exposure-related data in the literature and repositories, harmonising terms from different literature articles, and integrating across data sources. To characterize the existing exposure data landscape in Europe and highlight some of these challenges, we explored exposure-related data in the published literature (not restricted to Europe), existing European data repositories, and emerging data collection and production initiatives. In addition, relevant European policies were identified to broadly characterize regulatory requirements with respect to exposure data (see also Bruinen de Bruin et al., 2022). Through mapping existing exposure-related data to European regulatory data requirements, we develop an overview of chemical exposure and related health data (e.g., epidemiological data, human and animal toxicological data) across different branches of the European exposure data landscape and discuss the most prevalent challenges.

2.1. Mapping exposure data in the literature and European exposure-related data repositories

To identify literature contributing to the exposure data landscape, the European TIM Technology Editor (https://timanalytics.eu) was used to extract article titles and keywords from publications available in the source-neutral abstract and citation database Scopus (https://scopus.com). Studies were identified with the following set of chemical exposure-related terms present in the abstract, title, or author keywords: expos* AND (chemical* OR substance* OR pesticid* OR plant protect* OR pharmaceutic* OR medic* OR cosmetic* OR biocid* OR pollut* OR contamina* OR metal* OR compound*). The analysis included all articles from the period 2010 – spring 2022 (not restricted to European authors), resulting in titles and keywords for a total of 152,896 identified articles, demonstrating the massive amount of chemical exposure-related publications in the peer-reviewed literature. For our analysis, we consider literature keywords as a proxy for data that can be extracted from different research articles, either as supplementary data downloads or from pulling relevant data values from the text body.

Ignoring case differences (e.g., “Air pollution” vs “air pollution”), a total of 239,238 author keywords were associated with the identified articles. The keywords associated with the most articles include “air pollution” (4821 articles), “particulate matter” (2132 articles), “cadmium” (2705 articles), and “heavy metals” (2651 articles). For each of these top keywords, several variants of each term were also present. For example, there were 247 keywords that were variants of “air pollution”. While some variants were more common terms like “indoor air pollution” (307 articles), 182/247 terms were only keywords in individual articles, including “air pollution alert service”, “low-level air pollution”, and “winter air pollution”. The relationships between related top keywords (e.g., “air pollution” and “particulate matter”), their variants (e.g., “indoor air pollution”, “fine particulate matter”), and other keywords found in the same research articles (e.g., “asthma”) are shown in Fig. 1a-b.

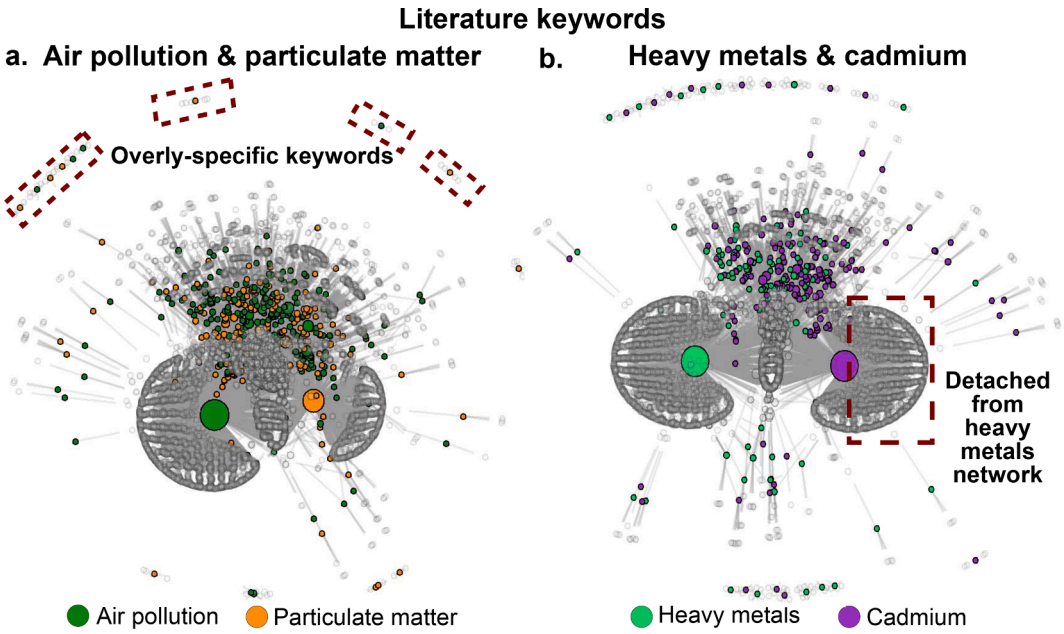
The sheer number of keywords present in the literature highlights the need for enhanced metadata and data harmonisation efforts, including alignment of terminology, consistent data associations, and clear definitions for different terms. If keywords are to represent the data available in a research article, then they should be concise, consistently used by authors, and reflective of the information contained within an article. Across all keywords curated from the exposure-related literature, 75 % were present in only one research article. This disparity is demonstrated in the small clusters of keywords completely separated from the main networks in Fig. 1a-b. The separation of some terms from the main network of exposure-related keywords also underscores the need to decrease separation between different facets of exposure science and enhance multidisciplinary research efforts. For example, while particulate matter is typically studied in the context of air pollution, several air pollution-related keywords were detached from the particulate matter network (Fig. 1a). This means that an individual searching for the term “particulate matter” among exposure-related information may miss the wealth of air pollution-related information unless they were also searching for “air pollution”. This is also true of heavy metals and cadmium (a heavy metal, Fig. 1b).

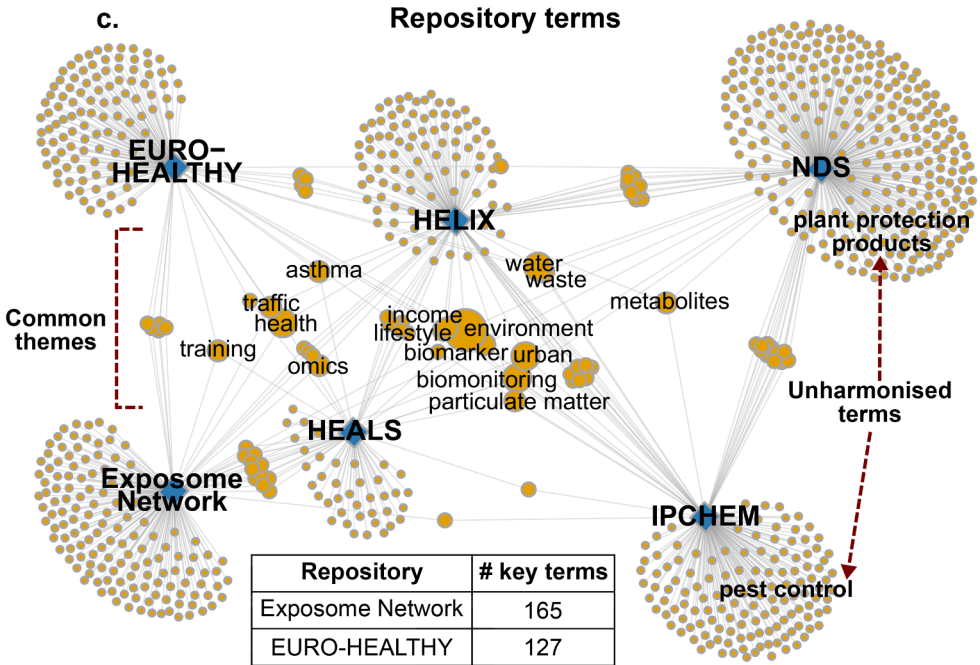
To characterize the exposure data landscape in existing data repositories, European experts from diverse international data repository initiatives described the data extraction, analytical tools, and content of the repository hosted at or provided by their respective institutions. An overview of these repositories is given in Table 1.

To characterize differences across data repositories and areas of overlap, a set of key terms was developed for three existing repositories (IPCHEM, NDS, EURO-HEALTHY) and three emerging data initiatives identified through ongoing European exposure data generation efforts (HELIX, Exposome Network, HEALS). Each term list was developed by collecting descriptors from data repository documents (e.g., Norman Database System’s glossary terms, EUROHEALTHY’s Atlas terms), “search terms” from existing data repository interfaces (excluding institution names, numeric ages and timeframes, and individual chemical identifiers), and focus areas/work package titles from emerging data generation initiatives. Each string of terms was broken down into concise keywords (e.g., HEALS work package title “Allergy and asthma – link with particulate matter (PM) and biologicals” was broken down into “Allergy”, “Asthma”, “Particulate matter”, “PM”, and “Biologicals”). A network of these key terms across repositories is in Fig. 1c.

Of the terms unique to each repository, different exposure themes can be identified. For example, with sampling media, IPCHEM includes several terms that identify specific sources where chemicals could be found (e.g., “clothes”, “cosmetics”, “appliances”), specific populations (e.g., “mothers”, “elderly”, “newborns”), and biomonitoring-related terms (e.g., “cord blood”, “urine”, “breast milk”). In contrast, NDS contains specific country and location names where monitoring occurs, chemical classes based on adverse outcome (e.g., “neurotoxicants”, “mycotoxins”), and different assay types and ecotoxicity threshold values for surface and marine waters, sediments, and biota. Emerging initiatives also cover different exposure-science areas, with HELIX covering more population behaviour (e.g., “physical activity”, “drinking”) and specific environment descriptors (e.g., “heating”, “sunlight exposure”), and the Exposome Network describing more general environment descriptors (e.g., “internal environment”, “external environment”) and data analysis techniques (“immunomics”, “metabolomics”, “risk assessment”).

Across repositories, the majority of key terms (91 %) were unique to each data resource or initiative, indicating each resource could meet different exposure data needs for a user. Of the terms that overlap, “environment” was the only term found in all databases, with “water”, “health”, “urban”, and “biomonitoring” found in four of the six. Other common themes include water and air descriptors (e.g., “surface water”, “indoor air”), health outcomes (e.g., “asthma”, “diabetes”), lifestyle descriptors (e.g., “alcohol consumption”), and stressor types (e.g., “Perfluoroalkyl chemicals (PFAS)”, “particulate matter”, “traffic noise”). However, as with the literature keywords, areas where a more consistent set of terms needs to be developed to harness available information across data repositories were also identified. For example, different resources may provide diverse information for the same chemical, but terms describing common stressor groups differed across repositories. NDS uses “plant protection products” and 16 sub-categories (e.g., “fungicides” and “insecticides”) to describe pesticides, while IPCHEM uses the term “pest control” (Fig. 1c). Terms for population descriptors and behaviours differed as well. IPCHEM contains “smoking” and “passive smoking”, EURO-HEALTHY contains the term “smokers”, and NDS contains “third-hand smoke” and “smoke”. This means that an





Repository	# key terms
Exposome Network	165
EURO-HEALTHY	127
HEALS	66
HELIX	133
IPCHEM	168
NDS	309

Fig. 1. Networks of author keywords in the literature and data repository terms. a) and b) Networks of most frequent author keywords (e.g., "air pollution" and "particulate matter"), variants (e.g., "indoor air pollution", "fine particulate matter"), and other associated keywords (e.g., "ozone", "asthma") from the exposure-related literature. Each node (circle) is an author keyword, and the grey lines between the nodes (edges) indicate at least one research article had those two keywords together. Coloured nodes = top keywords and variants in the literature, Grey nodes = other associated keywords. Connections between grey nodes are not shown. Node size is number of articles with each keyword ranging from 1 (smallest nodes) to 4821 (largest node). Dashed boxes correspond to highlighted challenges further described in the main text. c) Network of key terms between European data repositories and emerging data initiatives. Circle size = number of data repositories associated with that key term, range 1 (individual nodes on each database) – 6 (central "environment" labelled node). Terms labelled in bold represent areas for increased harmonisation between repositories, other labelled terms are present in at least 3 repositories (representing existing areas of overlap). The full number of key terms in each data repository is reported in the table below the figure. All labelled key terms are among literature keywords.

Table 1
Existing data repositories hosting European exposure-related data and emerging data-generating initiatives.

Data repository name	Hosting organization	Existing data repository or data-generation initiative	Types of exposure-related data/ information
IPCHEM - the Information Platform for Chemical Monitoring https://ipchem.jrc.ec.europa.eu	European CommissionJoint Research Centre (JRC)	Existing data repository	• Monitoring data (human, food, feed, environment, products, indoor air)
European Food Safety Authority: Data collection, standardization and analysis EFSA Scientific Data Warehouse EFSA knowledge junction (Zenodo) https://zenodo.org/communities/efsa-kj	European Food Safety Authority (EFSA)	Existing data repository	• Monitoring data (food) • Stressor properties (human toxicity) • Stressor use data (chemical occurrence, food consumption)
ECHA REACH database	European Chemicals Agency (ECHA)	Existing data repository	• Stressor properties (physiochemical properties, fate, human toxicity, ecotoxicity) • Stressor use data (exposure, emissions, use patterns, use tonnage bands)
NORMAN Database System (NDS) https://www.norman-network.com/nds	NORMAN Association	Existing data repository	• Monitoring data (environment, biota) • Stressor properties (physicochemical properties, ecotoxicity) • Stressor use data (use category)
World Environment Situation Room https://wesr.unep.org	United Nations Environment Programme (UNEP)	Existing data repository	• Monitoring data (environment, biota) • Receptor information (demography/ lifestyle, physical/ technical environment characteristics)
Exposome-Explorer https://exposome-explorer.iarc.fr	International Agency for Research on Cancer (IARC)	Existing data repository	• Monitoring data (human) • Receptor information (disease burden)
eChem Portal https://www.echemportal.org	Organisation for Economic Cooperation and Development (OECD)	Existing data repository	• Stressor properties (physicochemical properties, chemical classifications)
AirBase – Air Quality e-reporting https://www.eea.europa.eu/data-and-maps/data/aqereporting-9	European Environment Agency (EEA)	Existing data repository	• Monitoring data (air quality data)
WaterBase – Water Quality ICM https://www.eea.europa.eu/data-and-maps/data/waterbase-water-quality-icm-1	EEA	Existing data repository	• Monitoring data (water quality data)
The European Monitoring and Evaluation Programme (EMEP) https://www.emep.int	Convention on Long-range Transboundary Air Pollution (CLRTAP)	Existing data repository	• Monitoring data (environment)
EURO-HEALTHY (Shaping EUROpean policies to promote HEALTH equitY) https://healthyregionseurope.uc.pt	15 multidisciplinary institutions from 12 European countries	Existing data repository	• Receptor information (demography/ lifestyle, physical/ technical environment characteristics, disease burden)
Human Exposure Assessment Tools Database (heatDB) https://ecetoc.org/tools/ecetoc-heat-db	European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC)	Existing data repository	• Stressor tools (fate/ exposure, estimation tools)
EXPOsOMICS https://exposomics-project.eu	Thirteen of the world's leading organizations in the field	Data-generation initiative	• Monitoring data (air pollutants, water contaminants exposure data) linked to receptor information (biochemical and molecular changes via 'omics' techniques)
Health and Environment-wide Associations based on Large population Surveys (HEALS) https://cordis.europa.eu/project/id/603946	Sorbonne Université (France)	Data-generation initiative	• Monitoring data (Environmental, human exposure data) • Receptor information (socio-economic, biomarker and health effect data)
HELIX: The Human Early-Life Exposome https://www.projecthelix.eu	Instituto de Salud Global de Barcelona (ISGlobal)	Data-generation initiative	• Receptor information (exposome data for mothers and children from prospective birth cohort studies)
European Human Exposome Network https://www.humanexposome.eu	Nine research projects funded by Horizon 2020	Data-generation initiative*	• Monitoring data (human, environment) • Receptor information (demography/ lifestyle,

(continued on next page)

Table 1 (*continued*)

Data repository name	Hosting organization	Existing data repository or data-generation initiative	Types of exposure-related data/ information
			metagenomics and exposome data from varied population-based cohorts, environment characteristics, disease burden)
HazChem@Work	European Agency for Safety and Health at Work (EU-OSHA)	Data-generation initiative	• Monitoring data (occupational exposure measurements)

*Nine data generation initiatives – types of data/information condense the anticipated output across the 9 projects.

individual searching for smoking or pesticide-related information across different repositories would need to use different terms depending on the repository being searched. These examples represent areas for increased interoperability between exposure-related data repositories if key terms are harmonised.

2.2. European chemical policies requiring exposure data

To assess linkages between the exposure data found in the literature and data repositories and exposure data requirements in European policies, a set of exposure-related European regulations and directives was screened and themes across these policies and directives were characterized based on associated EuroVoc terms that describe their contents. Policies and directives were identified through searches conducted in the EUR-Lex database (https://eur-lex.europa.eu), by searching in policy documents (title and body text): expos* AND (chemical* OR substance* OR pesticid* OR plant protect* OR pharmaceutic* OR medic* OR cosmetic* OR biocid* OR pollut* OR contamina* OR metal* OR compound*) (see also Bruinen de Bruin et al. (2022) for further details). Resulting policy documents (directives and regulations), were analysed using a workflow developed in the KNIME platform (https://knime.com) that, using the described regular expressions, extracted the title, the URL and the associated EuroVoc descriptors of each policy document. EuroVoc terms are part of the multilingual, multidisciplinary EuroVoc thesaurus maintained by the EU to cover relevant activities (https://eur-lex.europa.eu/browse/eurovoc.html). EuroVoc terms are organized in 127 sub-domains. To ensure that the identified policies and directives were exposure-related rather than just mentioning a chemical briefly in the text body, the policies and directives were restricted to those with at least one EuroVoc descriptor clustered under exposure-relevant sub-domains: chemistry; iron, steel and other metal industries; means of agricultural production; plant product; agricultural activity; foodstuff; means of agricultural production; processed agricultural produce; agri-foodstuffs; electrical and nuclear industries; coal and mining industries; leather and textile industries; wood industry; and environmental policy. This identified 304 policies and 365 directives. Each of these policies and directives had some reference to exposure (based on the expos* search requirement) and had a corresponding set of EuroVoc terms (range of 2–12 terms, median of 6 terms). Across policies and directives, 664 unique EuroVoc descriptors were used (219 from the sub-domains). The list of policies/directives and corresponding EuroVoc descriptors are provided in SI Table S1.

Among the EuroVoc descriptors, "market approval", "marketing standard", "chemical product", "health risk", "foodstuff", "plant health control", and "food safety" were all present in over 100 policies/directives. Several agriculture-related (both crop and animal) and food-related (both food and packaging) terms were present in multiple policies/directives.

Of the 664 EuroVoc terms, 330 were unique to one policy/directive. These terms varied in character and included country names, element names (e.g., "iron", "carbon", "zinc"), specific plant types (e.g., "aromatic plant", "perennial vegetable", "citrus fruit") and other general terms and practices (e.g., "underground storage of waste", "bad weather", "zootechnics"), meaning the data requirements for these policies may need to be more targeted. This highlights the importance of exposure data generation directly in line with EU chemical policy needs.

As part of the Chemicals Strategy for Sustainability Towards a Toxic-Free Environment, the European Commission has established actions for 2020–2024 to simplify and consolidate the EU chemical legal framework to decrease overlaps and overcome inefficiencies in chemical assessment (European Commission, 2020b). Some of the actions that have identified particular needs or requirements for data harmonisation across policies are:

- The establishment of a 'One substance, one assessment' (1S1A) process to coordinate the hazard/risk assessment of chemicals across chemical legislation (European Commission, 2019a). Assessments carried out within different legislative domains often yield seemingly different outcomes and sometimes incoherent risk assessments, which can cause uncertainty for policy makers and increase assessment costs (e.g., the same chemical may be regulated differently under REACH regulation, the Food packaging regulation, and the Cosmetics regulation as it was assessed by different EU agencies/committees) (EFSA and ECHA, 2020). The 1S1A process is expected to increase efficiency for risk assessment and management of the same chemical for its marketing and authorization under different regulatory frameworks with specific timelines, data requirements, and methodologies. For successful implementation, risk assessment procedures for the freshwater environment need to be harmonised for the five main European chemical registration frameworks (medicines for human use (Directive, 2001/83/EC), veterinary medicines (Directive, 2001/82/EC), pesticides (Reg (EC) No 528/2012), biocides (Reg (EC) No 528/2012) and industrial chemicals (REACH, Reg (EC) No 1907/ 2006)) (van Dijk et al., 2021).
- The establishment of an EU repository of human and environmental health-based limit values to promote the reuse and harmonisation of health-based limit values among EU risk assessors and managers (European Commission, 2022a).
- The establishment of an Open Platform on Chemical Safety Data and tools to facilitate the sharing, access, and reuse of data and information coming from all sources (European Commission, 2019b).
- The horizontal proposal to remove legislative obstacles for reuse of data, streamline data flow across legislation, and extend open data and transparency principles from the EU food safety sector to other pieces of chemical legislation.

3. Strategic objectives for improving data repositories and analytics

The European Commission has increasingly included requirements in the Horizon Calls to streamline and integrate data generated by EU funded projects into data repositories that are sustained long-term. Therefore, it is essential that data be published following standard approaches. Our vision is to advance exposure data generation through improved data repositories and analytics, and to enhance exposure data usability and interoperability for increased uptake into European and global chemical policies and research on health and sustainability. However, numerous challenges inhibit cross-utilization between the literature, databases, and data platforms and therefore hinder the uptake of existing exposure-related data into these policies. These challenges are summarized in SI Table S2 and further detailed in the following

sections.

3.1. Developing a consistent exposure data terminology

As a first step in the analysis of the exposure data landscape in Europe, we identified keywords from the exposure-related literature (as a proxy for data contained within the articles), key terms describing European data repositories, and EuroVoc descriptors from European exposure-related policies and directives. We found that, across these branches of exposure-related data, different terms are employed to refer to common elements. These term differences range from inconsistencies in plurality or parts of speech (e.g., "smoker" vs "smokers", "smoker" vs "smoking"), to differences in the word or phrase used to describe a common element (e.g., "tobacco" vs "smoking", "individuals who smoke" vs "smokers"). Consequently, points of linkage across exposure-related resources are difficult to identify without applying specific text mining techniques. For example, in the literature 626 keywords were related to smoke or tobacco, ranging from the same terms in data repositories ("smoker", "smoking") to more specific terms only present in individual articles (e.g., "African American smokers", "Barriers to smoking cessation"). Additionally, the term "smoke" was not among the EuroVoc terms for exposure-related policies/directives identified. Instead, "tobacco" is a EuroVoc descriptor for Directive 2014/40/EU, which concerns the manufacture, presentation, and sale of tobacco and related products. While these common themes can be identified through tedious, manual processing, the use of different vocabularies for common themes inhibits high-throughput initiatives to form linkages across the exposure data landscape and makes it more challenging to determine where there are data gaps. Further, European policies/directives use a unique vocabulary decided upon and approved by the inter-institutional GIL EuroVoc Committee. While EuroVoc concepts were initially aligned to the National Library of Medicine Medical Subject Headings (MeSH) controlled vocabulary, this alignment has not been updated since 2017. Therefore, this vocabulary largely differs from that in the rapidly expanding literature and data repositories, which can impede efforts to identify underutilized exposure science data in regulatory contexts. With vocabularies across the literature, data repositories, and policies and regulations harmonised, additional challenges related to harvesting and harmonising data across studies can be more readily addressed.

In addition to differences in terms used to describe data repositories themselves, there are numerous differences in data identifiers within resources. There is an urgent need to establish a common European inventory of chemicals and their transformation products using unique identifiers. Different identifiers are used to describe chemicals (e.g., different chemical names, CAS (Chemical Abstracts Service) Registry Numbers, International Chemical Identifiers (InChI)) and health outcomes (e.g., disease ontology ID (DOID), MeSH), which can inhibit cross-utilization across resources, or incorporation of new data into existing data repositories. These differences are also present in the literature with different studies choosing to employ different vocabularies, and different scientific journals requiring different data identifiers for publication (e.g., chemical IUPAC (International Union of Pure and Applied Chemistry) names or CAS rules). Therefore, the development of a consistent, representative exposure terminology that is implemented within the literature, data repositories, and policies/directives is necessary to increase the uptake of existing exposure data into regulations. A suitable starting point toward the development of such a data-related terminology is the harmonisation of a glossary for exposure science (Heinemeyer et al., 2021). Other ontologies covering different chemicals and exposure sources include the Chemical Information Ontology (CHEMINF) (Hastings et al., 2011) and the Exposure Science Ontology (ExO) (Mattingly et al., 2012) while ontologies like the Experimental Factor Ontology (EFO) (Malone et al., 2010) can help to harmonise data collected across experiments.

3.2. Increased data quality, transparency, access, availability, reliability, and usability

Data transparency, access, and availability challenges are apparent within European exposure-related data repositories. Although access to information on chemicals has improved (e.g., through REACH or EFSA's open access platforms), most data and information on chemicals (tonnages, use categories, emissions, etc.) can be difficult to obtain from the producers. Information is often scattered, of variable quality, incomplete, not publicly available, and even inaccessible for risk assessors and managers. It is important to assess chemical risk during the entire life cycle of chemicals, materials, and products (including waste and recycled products), but this is challenging due to the lack of shared information on chemical use, overall chemical flows, and on the exposure to humans and the environment (Fantke et al., 2016). This is also the case for occupational data wherein authorities need data related to exposure sources, use, possible routes of workplace exposure, current and historical exposure levels, and most common risk management measures in place. The lack of such data can impede more comprehensive risk assessment strategies addressing combined exposures scenarios (PARC, 2020). Data confidentiality can also prevent wide utilization of relevant datasets, such as human biomonitoring data (Zare Jeddi et al., 2021). For example, IPCHEM contains human biomonitoring data from the HBM4EU initiative (Ganzleben et al., 2017), but most of these data are either not publicly available (in compliance with the General Data Protection Regulation requirements), or are only available as aggregated data without submitting for ethics board approval. Further safeguarding of data confidentiality is implemented in the context of REACH, biocides, and EFSA Transparency Regulations, which all have granted procedural rights to data submitters to request confidentiality protection of certain data categories prior to dissemination. This can ensure intellectual properties data and confidential information are not subject to unauthorized or unfair commercial use, and that data protection periods and data exclusivity rules are respected. However, data protection may come at the expense of data transparency (e.g., disseminating only aggregated data may hinder usability in certain analyses), and can inhibit adequate assessment of chemical safety and delay uptake of relevant chemical data into science and policy.

Data that are not made available with publication in the literature are often difficult to obtain, either from authors being unreachable or unwilling to make their data available, or from data being lost (Savage and Vickers, 2009; Vines et al., 2014). Several publishers have adopted policies requiring underlying data to be published alongside articles, including PLOS and Springer Nature. While these requirements have increased publication of Data Availability Statements (DAS) alongside scientific articles, an analysis of PLOS ONE articles found that the majority (80 %) of articles did not publish data in a repository, instead stating the data were located in the paper itself or shared as supplemental information (Federer et al., 2018). However, fragmented data locations, difficult to use and integrate data formats, varying data quality, and inconsistent availability can prevent uptake of the published literature into exposure-related data repositories and in chemical policy. For example, in addition to challenges with data being FAIR (Findable, Accessible, Interoperable and Reusable data), data repositories frequently have internal controls for data quality and/or standard approaches to data analysis, which might prevent incorporation of relevant data developed using different methodology or missing important metadata in the literature. The quality of data and information stored in existing exposure-related data repositories are often the responsibility of the data owner, and criteria for reliability and acceptance of exposure data and tools vary across EU chemical legislation (Bruinen de Bruin et al., 2022). As such, exposure-related data are missing a standardized assessment of data quality criteria to inform policies (Comero et al., 2020).

3.3. Enhanced data storage and related infrastructure

In the context of the European Exposure Strategy, building a cross-sector framework implementing enhanced data storage and related infrastructure to allow for exchange and innovative reuse of data—in a manner compliant with data protection legislation of widely-used data platforms (such as those in Table 1)—is of the utmost importance. This will upscale and accelerate scientific developments and enable informed and robust exposure and risk assessments in support of EU evidence-driven policy making.

To meet the needs of exposure-related data users, there needs to be a balanced combination of centralized data hubs for broadly usable exposure-related data and decentralized databases that focus on specific exposure aspects and/or meet specific application context requirements. The legislative framework for the governance of common European Data Spaces and related opportunities specified in the European Data Strategy (European Commission, 2020a) and European Open Science Cloud (EOSC) (European Commission, 2018b) provide opportunities for operationalizing collaborative frameworks and enhanced storage and infrastructure. EOSC, in line with the principle 'as open as possible, as closed as necessary', will facilitate discovery, sharing of, access to, and use/reuse of data and services by researchers. This represents one approach to streamline exposure-related data for research and policy-oriented purposes. Further, to make data available for better policy making, the European Commission has established the EU Open Data Portal, which makes the data produced and/or funded by the European Commission publicly available via a common gateway (https://data.europa.eu/euodp/en/data/). This portal provides access to an expanding range of data from EU institutions and other EU bodies via a standardized catalogue. This gives straightforward access to EU open data that have been systematically assessed in terms of quality by data providers (including environmental and health data) and a list of apps and web tools reusing these data. Another resource, Re3data (https://www.re3data.org), is a global registry of research data repositories maintained by different universities and research institutions. Through Re3data, disparate, heterogeneous data from different academic disciplines are made available to policy makers and researchers to increase access and improve usability of research data.

3.4. Increased automation in data management

By building systematic, automated approaches to identify, extract, and integrate new data immediately upon publication, exposure data repositories can stay updated without requiring manual curation (Aurisano and Fantke, 2022). This is essential to keep up with the pace of exposure data and ensure consistent data quality and constant policy relevance. To complement European goals for the Digital Decade and build a European approach to artificial intelligence (AI), the European Commission developed an AI strategy (European Commission, 2018a). This strategy aims to streamline AI research and proposes policy options for AI regulation to enable development and uptake of AI in the EU Digital Europe Programme (European Commission, 2021b).

AI can help screen available information from different sources on chemical properties, exposure, effects, production information, and use. From set criteria for identifying the most critical chemicals (e.g., that pose the highest risk), AI can then learn and build indicators to flag priority chemicals. AI can also potentially help with complex evaluations, such as to determine possible effects of combined exposure to chemicals and simulate exposure to chemicals for a multitude of possible exposure scenarios across different sectors and pathways (e.g., food, consumer goods, occupational exposures), and for different populations (Wittwehr et al., 2020).

Data structure, interoperability, and quality are essential to data value, especially in the context of AI deployment (Emara et al., 2022; European Commission, 2020a). For example, different databases and software applications like IPCHEM, NDS, and the International Uniform Chemical Information Database (IUCLID), include monitoring data, physicochemical properties, and (eco)toxicity data that could be relevant for decision makers and researchers. However, these data sources can be challenging to cross-utilize in automated efforts owing to different data structures, terminologies, or information fields that can be submitted as free text (Fantke et al., 2020a). Therefore, the application of standard compatible formats and open protocols for gathering and processing exposure-related data from diverse sources across chemical sectors and policy domains should be promoted. This can be done according to the FAIR principles for data, taking into account the decisions of chemical sector-specific authorities on which data can be used, how, and by whom (e.g., for scientific research purposes).

3.5. Increased data integration

In order to incorporate available exposure data from the literature and data repositories into European regulations and other decision-making frameworks, there needs to be enhanced data integration initiatives. Different repositories cover different data streams, including human/environmental monitoring, physicochemical properties, and chemical hazard data. By progressively linking and integrating fragmented chemical data collection efforts, one of the critical requirements for enhanced cooperation among existing regulatory frameworks will be fulfilled, as advocated by the 1S1A approach. In the long run this approach envisages increased data management efficiency, more consistent exposure and risk assessments, and improved robustness in assessments on common datasets of hazard and use and exposure assessment tools and methodologies (CEFIC, 2021). However, numerous challenges can hinder data integration efforts, including the aforementioned challenges of standard terminologies, data transparency, and data storage. Other challenges include cross-harmonisation of data and interoperability challenges. Data extracted from diverse sources are often of different data quality, developed for different purposes using different methods, covering different scales (e.g., different spatiotemporal scales), and available in different formats that are not always easily integrated (Aurisano and Fantke, 2022; Kosnik et al., 2022; Lapatas et al., 2015). Harmonisation of data reporting formats and standards across chemical legislative domains will improve interoperability among data repositories (Zare Jeddi et al., 2021). At an international level, harmonised test guidelines and templates for toxicity, physicochemical, fate, and behaviour properties have been developed by the OECD. The OECD Harmonised Templates (OHTs) are standard data formats for reporting information used for risk assessment of chemicals. While primarily used for studies to determine chemical properties or effects on human health and the environment, OHTs 301 to 306 target data and information on the use of and exposure to chemicals.

By using common identifiers/ontologies, building complete and structured data in appropriate formats for the underlying information, and providing well-structured and informed metadata, interoperable data structures from various existing and new exposure-related data repositories and literature studies can be promoted and cross-utilized in support of EU policies. As an example of the value of data integration, we analysed the sets of terms identified from the exposure-related literature, data repositories, and chemical policies and directives to identify common themes and areas for potential cross-utilization. As a first analysis, direct (exact word) key term matches were identified across exposure data repositories and the literature, and matched to EuroVoc descriptors (Fig. 2a).

Only half of the EuroVoc descriptors had a direct match to keywords collected in the exposure-related literature analysis (62 % of the exposure-relevant sub-domain EuroVoc terms). Of those EuroVoc terms that did not overlap, many referred to specific food items (e.g., "leguminous vegetable", "groundnut oil") or environmental policy terms (e.g., "protection of animals", "protection of plant life"). Between the 968 total data repository terms and 669 EuroVoc terms, the majority (98.3 %) did not overlap. However, some terms did provide direct relevance to

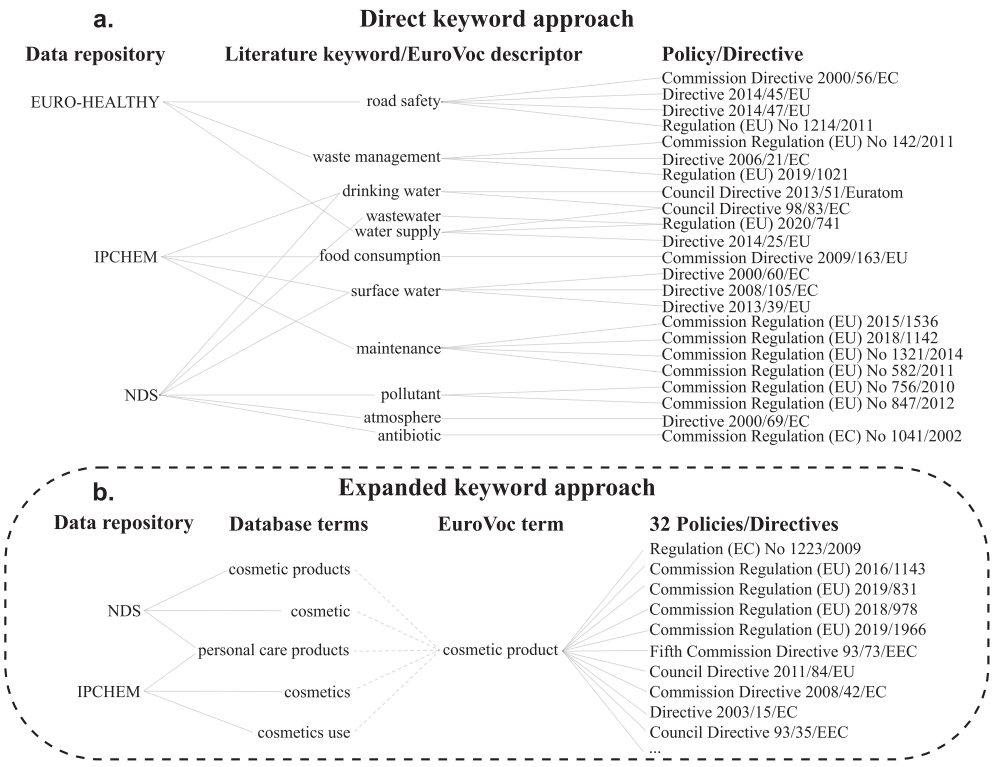


Fig. 2. Data repository key term matches with EuroVoc descriptors and corresponding policies and directives. a) Data repository terms that are exact-word matches with EuroVoc descriptors (labelled "Database term/EuroVoc term") are matched to policies. Using only direct term matches results in limited policy/directive matches per term. b) Example of possible expanded matches between data repositories and regulations. Multiple data repository terms (labelled "Database term") are harmonised with a "EuroVoc term", expanding possible applications for exposure-related data in 32 policies/directives. Note: all given database terms/EuroVoc terms match keywords identified in the exposure-related literature.

policy, and/or highlight areas for possible expansion in data utilization across EU regulations. For example, both IPCHEM and NDS had "surface water" as a key term for their data repositories, and "surface water" was a EuroVoc descriptor for Directives 2000/60/EC, 2008/105/EC, and 2013/39/EU, which are related to developing water policy frameworks and standards for quality and priority substances. Additionally, 126 literature articles had "surface water" as a keyword. Other policies/directives with direct matches between water key terms and EuroVoc descriptors include "drinking water" between IPCHEM, NDS, and Directives 98/83/EC and 2013/51/Euratom, and "wastewater" between NDS and Regulation (EU) 2020/741. This is in line with expectation, as NDS activities contribute to Directive 98/83/EC and Regulation (EU) 2020/741. Additionally, the key term "water supply" from EURO-HEALTHY and 12 research articles was a match for Regulation (EU) 2020/741 and Directive 98/83/EC, as well as Directive 2014/25/EU.

Repositories like IPCHEM and NDS are actively (directly or indirectly) contributing to exposure-related directives and regulations. However, not all areas of active synergy are currently captured. NDS contributes to Directive 2000/60/EC, Regulation (EU) No 528/2012, Regulation (EC) No 1107/2009, Directive 2010/75/EU, Directive 2008/56/EC, and Directive 2008/50/EC, among others. EuroVoc descriptors used by these legislations have themes matching those in NDS (e.g., EuroVoc descriptors include "atmospheric pollution", "prevention of pollution", and "plant health product" while NDS contains terms "atmosphere", "pollutant", and "plant protection products"), but no direct matches were formed. Similarly, IPCHEM has been acknowledged as the official reporting platform for monitoring data under Regulation (EU) 2019/1021 on persistent organic pollutants, yet there were no direct matches between generalized EuroVoc descriptors (e.g., "pollution control measures", "waste management") and more specific IPCHEM repository terms (e.g., IPCHEM lists specific monitoring locations and specific waste types). This is an identified area for improving interoperability of IPCHEM where one of the current objectives is to enhance data harmonisation for increased matching possibilities with various vocabularies when reporting chemical occurrences. Therefore, by establishing more standard vocabularies (e.g., data repositories develop descriptive terms for the data they contain based on EuroVoc descriptors) and/or using fuzzy matching techniques, more clusters of data themes matched between the literature, data repositories, and policies can be identified to enable more effective data integration. For example, clusters of EuroVoc descriptors can be formed around specific categories of exposure science data. The EuroVoc descriptor "cosmetic product" had 32 associated policies and directives, and by structuring looser key terms, can form matches with IPCHEM and NDS across key terms like "cosmetics use" and "personal care products" (Fig. 2b). By assessing potential information overlap between these data repositories and regulations, areas for increased exposure data uptake can be identified. In contrast, areas where there is no current overlap can be used as a starting point to identify data gaps in data repositories or the literature that need to be filled to better address policy needs.

3.6. Innovative tools for improved data retrieval, handling, and analysis

Development of systems and tools should be tailored to meet the requirements of different EU policies and regulations while facilitating broad application across the policy landscape. As part of the European strategy for exposure data, data-driven applications are promoted to increase the availability of and trust in data for policy and decision support (European Commission, 2020a). With increasing data volume, improved data analytics and processing is necessary to leverage this potential so new techniques can be efficiently incorporated and accepted into regulations as they are developed. A discussion of specific efforts from different agencies, repositories, and emerging data-generation initiatives (e.g., EFSA, IPCHEM, European Human Exposome Network) is given in SI Section S1.

Building interoperable data platforms requires innovative tools for data retrieval, handling, and integration, since data from different sources can be developed with varied methodologies, be of different quality, and be structured in difficult-to-use/join formats. Standardized templates for data collection and tools for data harmonisation and quality control are employed by different repositories to address these challenges and integrate diverse data sources. Additional tools can help

users sift through data, such as machine learning-based tools to scour the literature (van de Schoot et al., 2021), and analyse data, such as privacy-preserving tools to view individual-level data (e.g., DataSHIELD, https://www.datashield.org). There are also EU-wide initiatives to improve development and uptake of AI to advance data retrieval and handling (European Commission, 2018a).

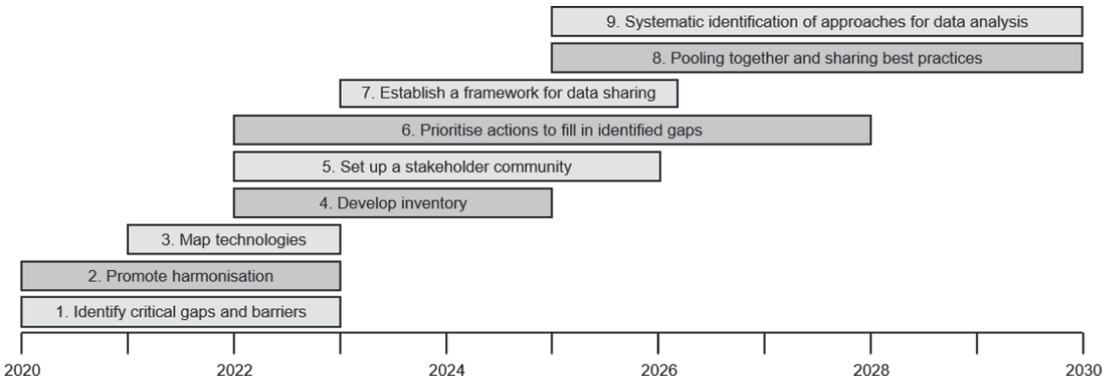
In addition to tools for data retrieval and integration, novel methods to analyse the wealth of diverse data are necessary to increase data use and knowledge generation. While many such tools exist for exposure-based risk assessment and modelling (Schlüter et al., 2022), they are not always freely publicly available, intuitive to use, or maintained long-term. Therefore, for a tool to be maximally employed, it should be continuously updated with clear version control mechanisms to track and uniquely refer to data, and old versions of the tool/data archived for retrieval. These tools should also have clear technical documentation and use guides interpretable for a variety of users, including agency and industry experts, researchers, and other stakeholders. For tools developed to meet specific regulatory needs, a major development underway by the European Commission is the development of the Common Open Data Platform on Chemicals, which aims to facilitate the sharing, access and reuse of information on chemicals coming from different sources. This platform will be achieved by building on existing consolidated and widely used tools and platforms (such as OHTs, IUCLID, QSAR (quantitative structure–activity relationship) Toolbox, IPCHEM, exposure models, and the Repository of Health Based Limit Values) and by streamlining chemical data flows and assessments managed by ECHA, EFSA, and EEA. This platform will aid cross-utilization and development of different tools and resources including common controlled vocabularies and chemical identifiers to support various end users (authorities, industry, academia, public) and prioritised use cases. It will also be maintained long-term through support by a sustainable EU IT infrastructure, an integral feature of the EU Green Deal Data Space (European Commission, 2022b).

4. Proposed action plan and recommendations for advancing exposure data and analytics in support of European science and policy

Considering the needs to foster exposure data uptake in policy as described in the previous sections, a set of specific actions are proposed to advance European exposure science and policy. An overview of these actions, how these feed into the overarching strategic objectives, relevant stakeholders we see contributing to each action, and a proposed timeline for implementation of these action items is provided in Fig. 3.

Following the timeline outlined by the ISES Europe working group, a systematic identification of data gaps in exposure-related data repositories has been initiated by relevant stakeholders in strict relation to relevant policy questions and associated use cases (Fig. 3, Action Item 1). For example, the monitoring data of the European Commission's Platform for Chemical Monitoring Data (IPCHEM) can support all steps in the regulatory cycle, including priorities for action, scientific evaluation, assessment of feasibility and impacts of a range of policy options, and subsequent evaluation of the impact and effectiveness of adopted polices/legislation. In a 2020 Workshop organized by the European Commission's Joint Research Centre, five use cases related to EU chemicals, environment, and health policy-related questions that IPCHEM data can contribute to when integrated with additional data were discussed (Franco et al., 2020). IPCHEM monitoring data have been used alone to support policies and the scientific community in Europe and globally (Kephalopoulos et al., 2019). However, combining IPCHEM monitoring data with other data types (e.g., toxicity data) to inform the aforementioned use cases revealed a number of challenges and limitations. These limitations differ on a case by case basis, including low availability of relevant data covering the required timeframes, exposure media, and geographic areas, lack of individual-level human biomonitoring and occupational exposure data owing to legal and ethical barriers for data sharing, discontinued monitoring efforts, and biases introduced by variability in substances monitored, sampling design, and analytical methods. Therefore, to advance the European Exposure Science Strategy, identification of data gaps in exposure-related repositories will be considered with respect to the specific objectives and questions to be addressed by chemical, environmental, and health policies and regulations. This will enable subsequent prioritization of actions to fill identified data-related gaps in line with needs of relevant stakeholders (Action Items 5 and 6).

Another short-term goal to improve exposure data generation and uptake in Europe is to increase interoperability of exposure data by promoting harmonised terms and vocabularies (Action Item 2). This has already been initiated by numerous stakeholders. The glossary for exposure science developed by ISES Europe can serve as a starting point for data-related terminology (Heinemeyer et al., 2021), but alignment of vocabularies across data repositories is also necessary. For example, the NORMAN Association considers StdInChIKeys to be the best alternative to the conventionally used chemical names and CAS registry Numbers. This work will also encourage journal requirements for data publication to follow more standard requirements for terminology and associated data/metadata (Vines et al., 2013). As keywords are used in the literature, in policies and regulations, and can inform data contained in data repositories, improved harmonisation across these terms is essential to increase data findability and usability. By encouraging authors to select keywords that can serve as metadata for the information contained within articles, literature can be more readily found for uptake into relevant repositories and used in policies and regulations. Alignment of EuroVoc and MesH terms will also be of great advantage to the research community for retrieval and utilisation of exposure and health-related data in support of various regulatory contexts. Further, maintenance of published exposure data is crucial, and numerous stakeholders have initiated this process (as a precursor to Action Item 7). For example, data storage and related infrastructure for exchange and reuse of exposure-related data has been initiated through the European Data Spaces, EOSC, and EU Open Data Portal as described in section 3.3. In Europe co-funded partnerships such as HBM4EU and the European Partnership for the Assessment of Risks from Chemicals (PARC) are promoted to support collaboration between European Commission Services, EU Agencies and national public organisations that conduct risk assessment and/or regulatory activities, and the contributing research communities. The PARC partnership is a joint research and innovation programme under development to strengthen the scientific basis for chemical risk assessment in the EU by bringing risk assessors and managers together with scientists to accelerate method development, data and knowledge generation, and to facilitate the transition to next generation evidence-based risk assessment. This meets many of the goals in our action plan. The PARC data collection will be interoperable with the main data repositories of EC Services and EU Agencies (IPCHEM and databases of ECHA, EFSA, and EEA) and PARC will develop tools and methods for combining and integrating heterogeneous data (e.g., environmental exposure, human biomonitoring, and different types of (eco)toxicity data) from available databases (PARC, 2020). Additionally, EU-OSHA started a pilot study in six EU Member States to collect data on occupational exposure to cancer risk factors (Riedmann et al., 2018), has published a data-driven methodology to assess exposure to dangerous substances in work places (Basinas et al., 2019), and collaborates with the JRC to integrate national occupational exposure measurement data into IPCHEM. Such wide partnerships between various stakeholders allow for more efficient use of existing resources (human, infrastructures, financial), knowledge, data, and best-practice transfer between countries and EU organizations. These partnerships also enable interconnections and interoperability of numerous data repositories in conjunction with integrated modelling capabilities for enhanced data storage, use of innovative ICT (information and communication technology) techniques, cloud computing, and related infrastructure (as needed for Action Items 3, 7, 9).



Action Item	Proposed actions, projects, tasks	Short title (in timeline)	Strategic Objective(s) (Section 3)	Global Objective(s)	Stakeholders
1	Identify ***critical data-related gaps and barriers to their use*** (e.g., constraints relating to spatial and temporal data availability and coverage, semantic and technical interoperability among distributed data and information platforms).	Identify critical gaps and barriers	Increased data transparency and availability	Reduce barriers to uptake of exposure science data in policy	Authorities, Practitioners, Data providers
2	Promote ***harmonisation of terms***, ***controlled vocabularies, templates for data and metadata documentation and reporting***, ***standardisation of data collection methods, enhanced data sharing,*** and ***development of indicators for data quality*** and comparability across exposure-relevant policy domains and regulations.	Promote harmonisation	Developing a consistent exposure data terminology, Enhanced data storage and related infrastructure	Increase interoperability of exposure-related data for application across regulations	Authorities, Practitioners, Data providers, Exposure scientists, Standardization bodies
3	Map the ***state of the art*** exploitation of ***technologies*** related to text mining and analysis, artificial intelligence, interconnection of big data domains, and visualisation tools ***with the aim of bridging the gap from data to knowledge.***	Map technologies	Increased data integration, Innovation for improved data analysis tools	Enhance usability of new methods and technologies for applying exposure science data to policy	Exposure scientists
4	Develop an ***inventory of existing data sources*** and ***already undertaken case studies*** that have been used in the EU and beyond linking chemicals, environmental data, exposure settings, and human data with specific health outcomes and considering environmental health inequalities.	Set up inventory	Enhanced data storage and related infrastructure, Increased data integration	Increase uptake of existing exposure science data into policy	Authorities, Practitioners, Data providers, Exposure scientists
5	***Set up a stakeholders community*** for exposure-related data repositories and analytics	Set up a stakeholder community		Reduce barriers to uptake of exposure science data in policy	ISES Europe
6	Prioritise ***actions to fill in the identified critical data-related gaps*** within respective timelines (short, medium, long-term) for defining and undertaking ***new case studies of high relevance for EU and MS policies and for advancing science***.	Prioritise actions to fill in identified gaps	Increased data integration	Reduce barriers to uptake of exposure science data in policy	Authorities, Funding bodies
7	Establish ***common principles and a framework for data sharing and use/reuse*** among stakeholders while respecting protection of intellectual property rights and personal (sensitive) data.	Establish a framework for data sharing	Increased data transparency and availability, Increased automation in data management	Increase uptake of new exposure science data into policy	Data providers, Exposure scientists, funding bodies
8	Pooling together and ***sharing best practices for data integration and interpretation, standardised and harmonised datasets*** and ***data formats*** used in connection with tools and models and ***performance indicators*** that are used in: policy making, exposure and health risk assessments, surveillance and monitoring activities, evaluation of intervention measures and health information systems.	Pooling together and sharing best practices	Enhanced data storage and related infrastructure, Increased data integration	Increase uptake of new exposure science data into policy	Practitioners, Data providers, Exposure scientists
9	Systematic identification, application and evaluation of ***novel methodologies, technologies, automatic, and semi-automatic approaches*** for data analysis	Systematic identification of approaches for data analysis	Innovation for improved data analysis tools	Enhance usability of new methods and technologies for applying exposure science data to policy	Authorities, Data providers, Exposure scientists, Funding bodies

Fig. 3. Proposed roadmap of actions in response to the strategic objectives for advancing exposure data and analytics. An estimated timeline of key actions with specific tasks related to the European Exposure Science Strategy—integrated into the global exposure science context—is provided.

An example of a large infrastructure for environmental data is the one proposed by the EU funded project SOLUTIONS (Brack, 2019) in collaboration with the NORMAN Network. This infrastructure has demonstrated the usefulness of a federated European infrastructure storing raw non-target screening data converted into a common, open format allowing for on-demand access to retrospective screening of high-resolution mass spectrometry (HR-MS) data. The parallel development of such infrastructure and a common European platform for harmonised, regularly-updated dataset information, including newly discovered environmental pollutants, will enable prioritization of pollutants that may cause adverse environmental effects at the local, national, and European scale in support of European water and chemicals management policies (Dulio et al., 2020; Slobodnik et al., 2019). However, it should be noted that these advances require not only conceptual and technical innovation in monitoring and assessment but may also demand better coordination of European chemical legislation (Kortenkamp et al., 2019), changes to the Water Framework Directive (European Commission, 2000), and an increase in cooperation, and cross-compliance of chemicals and water regulations (Brack et al., 2019). This highlights the importance of multi-stakeholder collaborations to meet these needs (Action Items 7, 9) and suggests that investment in establishing big data infrastructures and developing innovative tools for exposure and risk assessments can lead to scientifically rigorous outcomes supporting EU cross-cutting regulations.

In order to facilitate comparability and sharing of heterogeneous exposure-related data for use across various EU policies and regulations, a conceptual framework based on common definitions and standardised assessment of criteria and metrics to measure data quality should be developed. These quality indexes and scores should describe spatial, temporal, methodological and metrological traceability of data, and a good starting point could be the conceptual data framework developed for IPCHEM (Comero et al., 2020). With these advancements, the medium-term goals of ISES Europe for inventories of exposure science data and a framework for data sharing and reuse can be advanced, with stakeholder involvement and enforcement (Action Items 4, 5, and 7).

Long term, we recommend establishing an overarching framework and tiered approach based on fit-for-purpose methodologies, innovative tools, and best practices for exposure science to meet the needs of chemical, health, and environmental policies and regulations (Action Items 3, 8, and 9). This framework should target all relevant stakeholders who are involved in exposure data generation and use (policy makers and regulatory authorities, funding bodies, and groups producing and using chemicals including industry and general practitioners) and should:

(a) Define and prioritize relevant questions in the form of use cases aligned to actual regulatory needs and processes that could benefit from various types of data (e.g., monitoring data, health data, toxicity data, etc.), methodologies, and tools,

(b) Support coordinated allocation of the significant resources for and retrieval of multiple data streams and innovative tools (e.g., for exposure monitoring, integrated exposure and risk assessment, data mining, AI, etc.) that would be required by the use cases,

(c) Promote and facilitate linkage of data and models used across regulatory domains under a common scientific framework to consistently conduct and interlink exposure and risk assessments, which should be supported by methodologies and tools for scrutinizing available data to guarantee retrieval and use of high-quality data for specific policy and regulatory purposes,

(d) Coordinate the development of common data collection, reporting, and sharing procedures,

(e) Oversee the development, implementation, and assessment of measurable key performance and policy impact indicators and monitor EU funding schemes and national investments, and.

(f) Stimulate and streamline cooperation between Commission Services, EU agencies (ECHA, EFSA, EEA), EU Member States authorities and regulatory bodies, citizen science initiatives and organizations, industry, and the broader exposure science community.

This is of the utmost importance for building and progressively boosting a FAIR data culture enabling open science and enhancing exchange of data and information between different stakeholders and policy domains. In this context, interoperability and long-term sustainability of databases, hubs, and platforms is a key requirement that should be promoted and enforced. Further, to successfully operationalise this framework requires careful evaluation of the costs and trade-offs in standardisation and harmonisation and to identify a suitable level and structure of exposure data necessary to support exposure and risk assessment processes and decisions within and across regulations. In this context, we propose the framework be implemented:

(a) Continuing to push for harmonisation and standardisation of data formats, vocabularies, etc. This is essential to minimize efforts across regulations based on a clear “common” understanding of data characteristics and requirements (i.e. avoiding double efforts).

(b) Continuously reviewing and refining the decision contexts and related requirements for exposure data and their suitability to support these decisions.

(c) With continued learning from successful models of data repositories development and applications in other science fields to guide efforts for harmonising exposure data and databases.

5. Conclusions

To ensure that exposure science meaningfully informs EU policies and regulations, there is a need to align efforts for the generation, collection and use/reuse of exposure-related data to improve chemical risk assessment and management practices. Exposure-related data are essential for different decision contexts and building more consistent, transparent, and robust data production and management (based on current and future data analysis and digitalization trends) is necessary to improve uptake into science and policy. These needs align with the ongoing efforts of the European Exposure Science Strategy 2020–2030 (Fantke et al., 2022) and the needs of the European Chemicals Strategy for Sustainability as part of the European Green Deal. In this study, the existing exposure data landscape was mapped to requirements for data analytics and repositories across European policies and regulations, and needs and ways forward for improving data generation, sharing, and use were identified. We outline strategic objectives and propose an action plan to advance data repositories and analytics in exposure science. This includes the development of consistent exposure data standards and vocabularies for data production, collection, reporting, and analysis, increased data transparency and availability, enhanced data storage and related infrastructure, increased automation in data management, increased data integration, and advanced tools for innovative data analysis.

While there is still a lot of work to be done to achieve our goals, the present work of the ISES Europe Working Group on ‘Exposure data repositories and analytics’ is an important milestone to focus future research efforts and provide guidance to advance exposure data and their successful uptake in research and policy in Europe, in cooperation with key stakeholders and policy makers worldwide. To achieve these ambitious goals, international cooperation will be required to align and optimise funding opportunities (e.g., EU funded Horizon Programmes) and integrate efforts of European (e.g., the EU Green Deal, Circular Economy, European Strategy for Data) and international strategies (e.g., OECD Environmental Strategy, SAICM, UN Sustainable Development Goals) to achieve our overarching goals.

Disclaimer

The views and opinions expressed in this article are exclusively those of the authors and do not represent the official position of their respective institutions.

CRediT authorship contribution statement

Marissa B. Kosnik: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Visualization, Writing – original draft, Writing – review & editing. **Stylianos Kephalopoulos:** Conceptualization, Project administration, Writing – review & editing. **Amalia Muñoz:** Formal analysis, Writing – review & editing. **Nicolò Aurisano:** Writing – review & editing. **Alberto Cusinato:** Writing – review & editing. **Sani Dimitroulopoulou:** Writing – review & editing. **Jaroslav Slobodnik:** Writing – review & editing. **Jonathas De Mello:** Writing – review & editing. **Maryam Zare Jeddi:** Writing – review & editing. **Claudia Cascio:** Writing – review & editing. **Andreas Ahrens:** Writing – review & editing. **Yuri Bruinen de Bruin:** Writing – review & editing. **Lothar Lieck:** Writing – review & editing. **Peter Fantke:** Conceptualization, Methodology, Investigation, Supervision, Project administration, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The review of the manuscript by Dr Antonio Franco (EC, DG JRC) and Dr Alessandro Carletti (EC, DG JRC), is gratefully acknowledged. This study was executed by the ISES Europe working group on “Data repositories & analytics”, kicked-off during the 3rd European Exposure Science Strategy Workshop organised and supported by the European Commission’s Joint Research Centre on 30/31 Jan 2020 in Ispra, Italy (JRC project no 30032: DTHC – Digital Transformation of Health and Care). The study was financially supported by the “Safe and Efficient Chemistry by Design (SafeChem)” project funded by the Swedish Foundation for Strategic Environmental Research (grant no. DIA 2018/11), by the SPRINT project (grant no. 862568) funded under the European Union’s Horizon 2020 Research and Innovation program, and by the PARC project (grant no. 101057014) funded under the European Union’s Horizon Europe Research and Innovation program.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2022.107610.

References

Aurisano, N., Fantke, P., 2022. Semi-automated harmonization and selection of chemical data for risk and impact assessment. Chemosphere 302, 134886. https://doi.org/10.1016/j.chemosphere.2022.134886.

Backhaus, T., Scheringer, M., Wang, Z., 2018. Developing SAICM into a framework for the international governance of chemicals throughout their lifecycle: Looking beyond 2020. Integr. Environ. Assess. Manag. 14, 432–433. https://doi.org/10.1002/ieam.4052.

Basinas, I., Graveling, R., Dixon, K., Ritchie, P., 2019. Developing a data-driven method for assessing and monitoring exposure to dangerous substances in EU workplaces. Eur. Agency Saf. Health at Work (EU-OSHA). https://osha.europa.eu/sites/default/files/Developing_data_driven_method.pdf.

Brack, W., 2019. Solutions for present and future emerging pollutants in land and water resources management. Policy briefs summarizing scientific project results for decision makers. Environ. Sci. Eur. 31, 74. https://doi.org/10.1186/s12302-019-0252-7.

Brack, W., Ait-Aissa, S., Backhaus, T., Birk, S., Barceló, D., Burgess, R., Cousins, I., Dulio, V., Escher, B.I., Focks, A., van Gils, J., Ginebreda, A., Hering, D., Hewitt, L.M., Hilscherová, K., Hollender, J., Hollert, H., Köck, M., Kortenkamp, A., de Alda, M.L., Müller, C., Posthuma, L., Schüürmann, G., Schymanski, E., Segner, H., Sleeuwaert, F., Slobodnik, J., Teodorovic, I., Umbuzeiro, G., Voulvoulis, N., van Wezel, A., Altenburger, R., 2019. Strengthen the European collaborative environmental research to meet European policy goals for achieving a sustainable, non-toxic environment. Environ. Sci. Eur. 31, 63. https://doi.org/10.1186/s12302-019-0232-y.

Bruinen de Bruin, Y., Lequarre, A.-S., McCourt, J., Clevestig, P., Pigazzani, F., Zare Jeddi, M., Colosio, C., Goulart, M., 2020. Initial impacts of global risk mitigation measures taken during the combatting of the COVID-19 pandemic. Saf. Sci. 128, 104773 https://doi.org/10.1016/j.ssci.2020.104773.

Bruinen de Bruin, Y., Franco, A., Ahrens, A., Morris, A., Verhagen, H., Kephalopoulos, S., Dulio, V., Slobodnik, J., Sijm, D.T.H.M., Vermeire, T., Ito, T., Takaki, K., De Mello, J., Bessems, J., Zare Jeddi, M., Tanarro Gozalo, C., Pollard, K., McCourt, J., Fantke, P., 2022. Enhancing the use of exposure science across EU chemical policies as part of the European Exposure Science Strategy 2020–2030. J. Expo. Sci. Environ. Epidemiol. 32 (4), 513–525. https://doi.org/10.1038/s41370-021-00388-4.

Carney Almroth, B., Cornell, S.E., Diamond, M.L., de Wit, C.A., Fantke, P., Wang, Z., 2022. Understanding and addressing the planetary crisis of chemicals and plastics. One Earth 5, 1070–1074. https://doi.org/10.1016/j.oneear.2022.09.012.

CEFIC, 2021. Cefic views on the Open Data Platform.

Chu, J.S.G., Evans, J.A., 2021. Slowed canonical progress in large fields of science. Proc. Natl. Acad. Sci. 118 https://doi.org/10.1073/pnas.2021636118.

Comero, S., Dalla Costa, S., Cusinato, A., Korytar, P., Kephalopoulos, S., Bopp, S., Gawlik, B.M., 2020. A conceptual data quality framework for IPCHEM – The European Commission Information Platform for chemical monitoring. TrAC Trends Anal. Chem. 127, 115879 https://doi.org/10.1016/j.trac.2020.115879.

Dorne, J.L.C.M., Richardson, J., Livaniou, A., Carnesecchi, E., Ceriani, L., Baldin, R., Kovarich, S., Pavan, M., Saouter, E., Biganzoli, F., Pasinato, L., Zare Jeddi, M., Robinson, T.P., Kass, G.E.N., Liem, A.K.D., Toropov, A.A., Toropova, A.P., Yang, C., Tarkhov, A., Georgiadis, N., Di Nicola, M.R., Mostrag, A., Verhagen, H., Roncaglioni, A., Benfenati, E., Bassan, A., 2021. EFSA’s OpenFoodTox: An open source toxicological database on chemicals in food and feed and its future developments. Environ. Int. 146, 106293 https://doi.org/10.1016/j.envint.2020.106293.

Dulio, V., Koschorreck, J., van Bavel, B., van den Brink, P., Hollender, J., Munthe, J., Schlabach, M., Aalizadeh, R., Agerstrand, M., Ahrens, L., Allan, I., Alygizakis, N., Barcelo’, D., Bohlin-Nizzetto, P., Boutroup, S., Brack, W., Bressy, A., Christensen, J. H., Cirka, L., Covaci, A., Derksen, A., Deviller, G., Dingemans, M.M.L., Engwall, M., Fatta-Kassinos, D., Gago-Ferrero, P., Hernández, F., Herzke, D., Hilscherová, K., Hollert, H., Junghans, M., Kasprzyk-Hordern, B., Keiter, S., Kools, S.A.E., Kruve, A., Lambropoulou, D., Lamoree, M., Leonards, P., Lopez, B., López de Alda, M., Lundy, L., Makovinská, J., Marigómez, I., Martin, J.W., McHugh, B., Miège, C., O’Toole, S., Perkola, N., Polesello, S., Posthuma, L., Rodriguez-Mozaz, S., Roessink, I., Rostkowski, P., Ruedel, H., Samanipour, S., Schulze, T., Schymanski, E. L., Sengl, M., Tarábek, P., Ten Hulscher, D., Thomaidis, N., Togola, A., Valsecchi, S., van Leeuwen, S., von der Ohe, P., Vorkamp, K., Vrana, B., Slobodnik, J., 2020. The NORMAN Association and the European Partnership for Chemicals Risk Assessment (PARC): let’s cooperate! Environ. Sci. Eur. 32 (1) https://doi.org/10.1186/s12302-020-00375-w.

EFSA, 2011. Use of the EFSA comprehensive european food consumption database in exposure assessment. EFSA J. 9 https://doi.org/10.2903/j.efsa.2011.2097.

EFSA, ECHA, 2020. In support of the EU chemicals strategy for sustainability: One substance – one assessment.

Emara, Y., Koroušić Seljak, B., Gibney, E.R., Popovski, G., Pravst, I., Fantke, P., 2022. Workflow for building interoperable food and nutrition security (FNS) data platforms. Trends Food Sci. Technol. 123, 310–321. https://doi.org/10.1016/j.tifs.2022.03.022.

European Commission, 2006. Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). Brussels, Belgium.

European Commission, 2008. Directive 2008/98/EC of the European Parliament and of the Council of 19 November 2008 on waste and repealing certain Directives. In: Commission E, (ed). Brussels, Belgium.

European Commission, 2015. Communication from the Commission - Closing the loop - An EU action plan for the Circular Economy. COM(2015) 614 final. Brussels, Belgium.

European Commission, 2016. Service contract to create a database and develop a model to estimate the occupational exposure for a list of hazardous chemicals in the Member States of the European Union and in the EFTA / EEA countries . Directorate B. Publications Office of the European Union, Luxembourg. 10.2767/863737.

European Commission, 2017. Study for the strategy for a non-toxic environment of the 7th Environment Action Programme: Final Report. Brussels, Belgium.

European Commission, 2018a. COM/2018/237 final - Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Artificial Intelligence for Europe. Brussels, Belgium.

European Commission, 2018b. Commission Staff Working Document. Implementation Roadmap for the European Open Science Cloud. SWD (2018) 83 final. Brussels, Belgium.

European Commission, 2019a. COM/2019/640 final - Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: The European Green Deal. Brussels, Belgium.

European Commission, 2019b. ENV/2019/OP/0014 Feasibility Study on a Common Open Platform on Chemical Safety Data.

European Commission, 2020a. COM/2020/66 final - Communication from the Commission to the European Parliament, the European Council, the Council, the

European Economic and Social Committee and the Committee of the Regions: A European strategy for data. Brussels, Belgium.
European Commission, 2020b. COM/2020/667 final - Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Chemicals Strategy for Sustainability Towards a Toxic-Free E. Brussels, Belgium.
European Commission, 2021a. Europe's Beating Cancer Plan: Implementation Roadmap.
European Commission, 2021b. COM/2021/118 final - Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: 2030 Digital Compass: the European way for the Digital Decad. Brussels, Belgium.
European Commission, 2022a. ENV/2021/OP/0019 Designing EU Repository of Health Based Limit Values and Collating Information for the First Version of the Repository.
European Commission, 2022b. Feasibility study on a common open platform on chemical safety data. Final report. Directorate General for the Environment. Brussels, Belgium.
Fantke, P., Ernstoff, A.S., Huang, L., Csiszar, S.A., Jolliet, O., 2016. Coupled near-field and far-field exposure assessment framework for chemicals in consumer products. Environ. Int. 94, 508–518. https://doi.org/10.1016/j.envint.2016.06.010.
European Commission, 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. Off. J. L 327, 0001–0073.
Fantke, P., Aurisano, N., Provoost, J., Karamertzanis, P.G., Hauschild, M., 2020a. Toward effective use of REACH data for science and policy. Environ. Int. 135, 105336 https://doi.org/10.1016/j.envint.2019.105336.
Fantke, P., von Goetz, N., Schlüter, U., Bessems, J., Connolly, A., Dudzina, T., Ahrens, A., Bridges, J., Coggins, M.A., Conrad, A., Hänninen, O., Heinemeyer, G., Kephalopoulos, S., McLachlan, M., Meijster, T., Poulsen, V., Rother, D., Vermeire, T., Viegas, S., Vlaanderen, J., Jeddi, M.Z., Bruinen de Bruin, Y., 2020b. Building a European exposure science strategy. J. Expo. Sci. Environ. Epidemiol. 30, 917–924. https://doi.org/10.1038/s41370-019-0193-7.
Fantke, P., Cinquemani, C., Yaseneva, P., De Mello, J., Schwabe, H., Ebeling, B., Lapkin, A.A., 2021. Transition to sustainable chemistry through digitalization. Chem 7, 2866–2882. https://doi.org/10.1016/j.chempr.2021.09.012.
Fantke, P., Bruinen de Bruin, Y., Schlüter, U., Connolly, A., Bessems, J., Kephalopoulos, S., Zare Jeddi, M., van Nieuwenhuyse, A., Dudzina, T., Scheepers, P. T.J., von Goetz, N., 2022. The European exposure science strategy 2020–2030. Environ. Int. 170, 107555 https://doi.org/10.1016/j.envint.2022.107555.
Federer, L.M., Belter, C.W., Joubert, D.J., Livinski, A., Lu, Y.-L., Snyders, L.N., Thompson, H., 2018. Data sharing in PLOS ONE: An analysis of data availability statements. PLoS ONE 13, e0194768.
Franco, A., Cusinato, A., Kienzler, A., Madia, F.A., Olvedy, M., Worth, A., Bopp, S., 2020. EU chemicals, environment and health policies: how can IPCHEM contribute? JRC Technical Report. JRC122907.
Ganzleben, C., Antignac, J.-P., Barouki, R., Castaño, A., Fiddicke, U., Klánová, J., Lebret, E., Olea, N., Sarigiannis, D., Schoeters, G.R., Sepai, O., Tolonen, H., Kolossa-Gehring, M., 2017. Human biomonitoring as a tool to support chemicals regulation in the European Union. Int. J. Hyg. Environ. Health 220, 94–97. https://doi.org/10.1016/j.ijheh.2017.01.007.
Groh, K., vom Berg, C., Schirmer, K., Tlili, A., 2022. Anthropogenic chemicals as underestimated drivers of biodiversity loss: scientific and societal implications. Environ. Sci. Technol. 56, 707–710. https://doi.org/10.1021/acs.est.1c08399.
Hastings, J., Chepelev, L., Willighagen, E., Adams, N., Steinbeck, C., Dumontier, M., 2011. The chemical information ontology: provenance and disambiguation for chemical data on the biological semantic web. PLoS ONE 6, e25513.
Heinemeyer, G., Connolly, A., von Goetz, N., Bessems, J., Bruinen de Bruin, Y., Coggins, M.A., Fantke, P., Galea, K.S., Gerding, J., Hader, J.D., Heussen, H., Kephalopoulos, S., McCourt, J., Scheepers, P.T.J., Schlueter, U., van Tongeren, M., Viegas, S., Zare Jeddi, M., Vermeire, T., 2021. Towards further harmonization of a glossary for exposure science—an ISES Europe statement. J. Expo. Sci. Environ. Epidemiol. https://doi.org/10.1038/s41370-021-00390-w.
Johnson, A.C., Jin, X., Nakada, N., Sumpter, J.P., 2020. Learning from the past and considering the future of chemicals in the environment. Science (80-.). 367, 384–387. https://doi.org/10.1126/science.aay6637.
Joint Research Center, 2020. IPCHEM - The European Commission's information platform for chemical monitoring data. Ispra, Italy.
Jolliet, O., Huang, L., Hou, P., Fantke, P., 2021. High throughput risk and impact screening of chemicals in consumer products. Risk Anal. 41, 627–644. https://doi.org/10.1111/risa.13604.
Kephalopoulos, S., Cusinato, A., Costa, S.D., Ceridono, M., Korytar, P., Bopp, S., 2019. Information Platform for Chemical Monitoring (IPCHEM)- Quotations. Impact, Recognition.
Kephalopoulos, S., Bopp, S.K., Costa, S.D., Cusinato, A., Lipsa, D., Geiss, O., 2020. Indoor air monitoring: Sharing and accessing data via the Information Platform for chemical monitoring (IPCHEM). Int. J. Hyg. Environ. Health 227, 113515. https://doi.org/10.1016/j.ijheh.2020.113515.
Kortenkamp, A., Faust, M., Backhaus, T., Altenburger, R., Scholze, M., Müller, C., Ermler, S., Posthuma, L., Brack, W., 2019. Mixture risks threaten water quality: the European Collaborative Project SOLUTIONS recommends changes to the WFD and better coordination across all pieces of European chemicals legislation to improve protection from exposure of the aquatic environment to. Environ. Sci. Eur. 31, 69. https://doi.org/10.1186/s12302-019-0245-6.
Kosnik, M.B., Hauschild, M.Z., Fantke, P., 2022. Toward assessing absolute environmental sustainability of chemical pollution. Environ. Sci. Technol. 56, 4776–4787. https://doi.org/10.1021/acs.est.1c06098.
Kristiansson, E., Coria, J., Gunnarsson, L., Gustavsson, M., 2021. Does the scientific knowledge reflect the chemical diversity of environmental pollution? – A twenty-year perspective. Environ. Sci. Policy 126, 90–98. https://doi.org/10.1016/j.envsci.2021.09.007.
Landrigan, P.J., Fuller, R., Acosta, N.J.R., Adeyi, O., Arnold, R., Basu, N. (Nil), Baldé, A. B., Bertollini, R., Bose-O'Reilly, S., Boufford, J.I., Breysse, P.N., Chiles, T., Mahidol, C., Coll-Seck, A.M., Cropper, M.L., Fobil, J., Fuster, V., Greenstone, M., Haines, A., Hanrahan, D., Hunter, D., Khare, M., Krupnick, A., Lanphear, B., Lohani, B., Martin, K., Mathiasen, K. V, McTeer, M.A., Murray, C.J.L., Ndahimananjara, J.D., Perera, F., Potočnik, J., Preker, A.S., Ramesh, J., Rockström, J., Salinas, C., Samson, L.D., Sandilya, K., Sly, P.D., Smith, K.R., Steiner, A., Stewart, R.B., Suk, W.A., van Schayck, O.C.P., Yadama, G.N., Yumkella, K., Zhong, M., 2018. The Lancet Commission on pollution and health. Lancet 391, 462–512. 10.1016/S0140-6736(17)32345-0.
Lapatas, V., Stefanidakis, M., Jimenez, R.C., Via, A., Schneider, M.V., 2015. Data integration in biological research: an overview. J. Biol. Res. 22, 9. https://doi.org/10.1186/s40709-015-0032-5.
Liu, Q., Li, L., Zhang, X., Saini, A., Li, W., Hung, H., Hao, C., Li, K., Lee, P., Wentzell, J.J. B., Huo, C., Li, S.-M., Harner, T., Liggio, J., 2021. Uncovering global-scale risks from commercial chemicals in air. Nature 600, 456–461. https://doi.org/10.1038/s41586-021-04134-6.
Malone, J., Holloway, E., Adamusiak, T., Kapushesky, M., Zheng, J., Kolesnikov, N., Zhukova, A., Brazma, A., Parkinson, H., 2010. Modeling sample variables with an experimental factor ontology. Bioinformatics 26, 1112–1118. https://doi.org/10.1093/bioinformatics/btq099.
Mattingly, C.J., McKone, T.E., Callahan, M.A., Blake, J.A., Hubal, E.A.C., 2012. Providing the missing link: the exposure science ontology ExO. Environ. Sci. Technol. 46, 3046–3053. https://doi.org/10.1021/es2033857.
Müller, N., de Zwart, D., Hauschild, M., Kijko, G., Fantke, P., 2017. Exploring REACH as a potential data source for characterizing ecotoxicity in life cycle assessment. Environ. Toxicol. Chem. 36, 492–500. https://doi.org/10.1002/etc.3542.
PARC, 2020. Draft proposal for a European Partnership under Horizon Europe - Partnership for the Assessment of Risk from Chemicals (PARC).
Riedmann, A., Banduch, I., Stepa-Aurora, R., Cavet, M., 2018. Feasibility study on the development of a computer-assisted telephone survey to estimate workers' exposure to carcinogens in the European Union. Luxembourg. https://doi.org/10.2802/390898.
SAICM, 2020. Strategic Approach to International Chemicals Management. United Nations Environment Programme.
Savage, C.J., Vickers, A.J., 2009. Empirical study of data sharing by authors publishing in PLoS Journals. PLoS ONE 4, e7078.
Schlüter, U., Meyer, J., Ahrens, A., Borghi, F., Clerc, F., Delmaar, C., Guardo, A., Dudzina, T., Fantke, P., Fransman, W., Hahn, S., Heussen, H., Jung, C., Koivisto, J., Koppisch, D., Paini, A., Savic, N., Spinazzè, A., Jeddi, M.Z., Goetz, N., 2022. Exposure modelling in Europe : how to pave the road for the future as part of the European Exposure Science Strategy. 10.1038/s41370-022-00455-4.
Slobodnik, J., Hollender, J., Schulze, T., Schymanski, E.L., Brack, W., 2019. Establish data infrastructure to compile and exchange environmental screening data on a European scale. Environ. Sci. Eur. 31, 65. https://doi.org/10.1186/s12302-019-0237-6.
United Nations Environment Programme, 2019a. Global Chemicals Outlook II - From Legacies to Innovative Solutions: Implementing the 2030 Agenda for Sustainable Development. Nairobi, Kenya. Report no.: ISBN: 978-92-807-3745-5.
United Nations Environment Programme, 2019b. UNEP circularity platform.
van de Schoot, R., de Bruin, J., Schram, R., Zahedi, P., de Boer, J., Weijdema, F., Kramer, B., Huijts, M., Hoogerwerf, M., Ferdinands, G., Harkema, A., Willemsen, J., Ma, Y., Fang, Q., Hindriks, S., Tummers, L., Oberski, D.L., 2021. An open source machine learning framework for efficient and transparent systematic reviews. Nat. Mach. Intell. 3, 125–133. https://doi.org/10.1038/s42256-020-00287-7.
van Dijk, J., Gustavsson, M., Dekker, S.C., van Wezel, A.P., 2021. Towards 'one substance – one assessment': An analysis of EU chemical registration and aquatic risk assessment frameworks. J. Environ. Manage. 280, 111692 https://doi.org/10.1016/j.jenvman.2020.111692.
Vines, T.H., Andrew, R.L., Bock, D.G., Franklin, M.T., Gilbert, K.J., Kane, N.C., Moore, J., Moyers, B.T., Renaut, S., Rennison, D.J., Veen, T., Yeaman, S., 2013. Mandated data archiving greatly improves access to research data. FASEB J. 27, 1304–1308. https://doi.org/10.1096/fj.12-218164.
Vines, T.H., Albert, A.Y.K., Andrew, R.L., Débarre, F., Bock, D.G., Franklin, M.T., Gilbert, K.J., Moore, J.-S., Renaut, S., Rennison, D.J., 2014. The availability of research data declines rapidly with article age. Curr. Biol. 24, 94–97. https://doi.org/10.1016/j.cub.2013.11.014.
Wittwehr, C., Blomstedt, P., Gosling, J.P., Peltola, T., Raffael, B., Richarz, A.-N., Sienkiewicz, M., Whaley, P., Worth, A., Whelan, M., 2020. Artificial Intelligence for chemical risk assessment. Comput. Toxicol. 13, 100114 https://doi.org/10.1016/j.comtox.2019.100114.
Zare Jeddi, M., Virgolino, A., Fantke, P., Hopf, N.B., Galea, K.S., Remy, S., Viegas, S., Mustieles, V., Fernandez, M.F., von Goetz, N., Vicente, J.L., Slobodnik, J., Rambaud, L., Denys, S., St-Amand, A., Nakayama, S.F., Santonen, T., Barouki, R., Pasanen-Kase, R., Mol, H.G.J., Vermeire, T., Jones, K., Silva, M.J., Louro, H., van der Voet, H., Duca, R.-C., Verhagen, H., Canova, C., van Klaveren, J., Kolossa-Gehring, M., Bessems, J., 2021. A human biomonitoring (HBM) Global Registry Framework: Further advancement of HBM research following the FAIR principles. Int. J. Hyg. Environ. Health 238, 113826. https://doi.org/10.1016/j.ijheh.2021.113826.