CellPress

One Earth

Perspective

Integrated data-driven cross-disciplinary framework to prevent chemical water pollution

Mohamed Ateia,[1,2,*] Gabriel Sigmund,[3,4] Michael J. Bentel,[5] John W. Washington,[6] Adelene Lai,[7,8] Nathaniel H. Merrill,[9] and Zhanyun Wang[10,*]

[1]United States Environmental Protection Agency, Center for Environmental Solutions & Emergency Response, Cincinnati, OH 45220, USA
[2]Department of Chemical and Biomolecular Engineering, Rice University, Houston, TX, USA
[3]Environmental Geosciences, Centre for Microbiology and Environmental Systems Science, University of Vienna, Josef-Holaubeck-Platz 2, 1090 Vienna, Austria
[4]Environmental Technology, Wageningen University & Research, P.O. Box 17, 6700 AA Wageningen, the Netherlands
[5]Department of Environmental Engineering and Earth Sciences, Clemson University, Clemson, SC 29634, USA
[6]United States Environmental Protection Agency, Center for Environmental Measurement and Modeling, Athens, GA 30605, USA
[7]Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, 6 Avenue du Swing, 4367 Belvaux, Luxembourg
[8]Institute for Inorganic and Analytical Chemistry, Friedrich-Schiller-University, 07743 Jena, Germany
[9]United States Environmental Protection Agency, Center for Environmental Measurement and Modeling, Narragansett, RI, USA
[10]Empa – Swiss Federal Laboratories for Materials Science and Technology, Technology and Society Laboratory, 9014 St. Gallen, Switzerland
*Correspondence: ibrahim.mohamed@epa.gov (M.A.), zhanyun.wang@empa.ch (Z.W.)
https://doi.org/10.1016/j.oneear.2023.07.001

SCIENCE FOR SOCIETY The world is facing formidable challenges of ubiquitous water pollution by toxic chemicals, which adversely impact human health and the environment with astronomical externalized costs on society. This perspective learns from past experience and proposes a novel, data-driven framework to break the current silos across different disciplines along the entire chemical life cycle. This framework can be applied to foster proactive action to address ongoing water pollution in a more effective, efficient, and timely manner by sharing relevant information among chemical production and use and environmental engineers. It can also be applied to prevent new water pollution by enabling better design of new chemicals and processes. To operationalize the framework, concerted efforts are needed to break existing silos and bring different disciplines together.

SUMMARY

Access to a clean and healthy environment is a human right and a prerequisite for maintaining a sustainable ecosystem. Experts across domains along the chemical life cycle have traditionally operated in isolation, leading to limited connectivity between upstream chemical innovation to downstream development of water-treatment technologies. This fragmented and historically reactive approach to managing emerging contaminants has resulted in significant externalized societal costs. Herein, we propose an integrated data-driven framework to foster proactive action across domains to effectively address chemical water pollution. By implementing this integrated framework, it will not only enhance the capabilities of experts in their respective fields but also create opportunities for novel approaches that yield co-benefits across multiple domains. To successfully operationalize the integrated framework, several concerted efforts are warranted, including adopting open and FAIR (findable, accessible, interoperable, and reusable) data practices, developing common knowledge bases/platforms, and staying vigilant against new substance "properties" of concern.

INTRODUCTION: THE HISTORICAL GENESIS OF CHEMICAL WATER POLLUTION

Clean water is a fundamental human need and the basis for a healthy ecosystem. However, the world is facing formidable challenges of water pollution by toxic chemicals.[1,2] For example, Andrews and Naidenko[3] estimated that in the United States alone, 18–80 million people may receive tap water contaminated by per- and polyfluoroalkyl substances (PFASs) at levels that may cause long-term health risks. Regarding the impact on ecosystems, a 2022 study found 26% of rivers (n = 258) across 104 countries contained at least one pharmaceutical pollutant exceeding safe levels for aquatic organisms.[4] In fact,

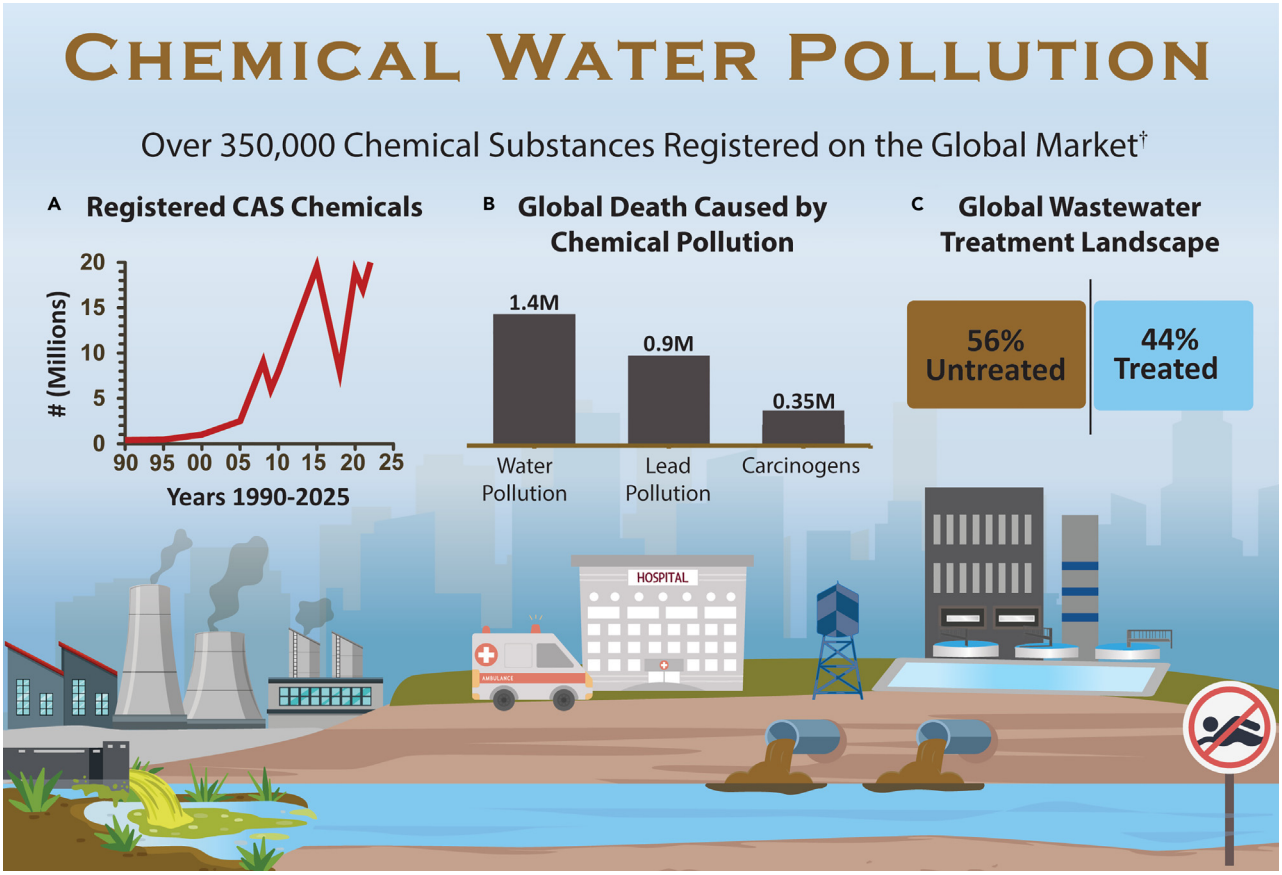

Figure 1. The challenges of chemical water pollution

(A) Since 1990, the number of chemicals registered by the Chemical Abstracts Service (CAS) per year has reached around 20 million chemicals in 2023. Data are from CAS.[8]

(B) Global deaths caused by chemical pollution, as modeled by Fuller et al.[9]

(C) The majority of global industrial and municipal wastewater streams remain untreated. Data are from Alabaster et al.[10]

†Data are from Wang et al.[11]

pollution—including water pollution—along with climate change and biodiversity loss has been recognized by the United Nations Environmental Program as "three self-inflicted planetary crises that are closely interconnected and put the well-being of current and future generations at unacceptable risk."[5]

Water pollution by toxic chemicals primarily originates from anthropogenic activities, both industrial and domestic, with some contribution from natural activities (e.g., naturally elevated arsenic in groundwater under reducing conditions).[6] In other words, during the acquisition, production, use, and disposal of chemicals and related products, many compounds have been and are being released from both point sources (e.g., wastewater spillage) and diffuse sources (e.g., road runoff).[7] In an ideal world, such releases would have been foreseen, with corresponding action taken in a timely manner to proactively minimize adverse impacts on human health and the environment (Figure 1). However, in reality, there has been a limited connection between each life cycle stage, from the upstream chemical innovation to the downstream development of water-treatment technologies.

The 20th century brought a drastic transformation of the global chemical industry. Discovery of new chemicals and their utility shifted from serendipity (e.g., Fleming's discovery of penicillin in 1928 and Plunkett's discovery of polytetrafluoroethylene [PTFE] in 1938) to utilizing novel cheminformatic systems capable of modeling interactions of multiple molecular functional groups with specific targeted systems and designing chemicals bearing desired properties. This shift has led to the continuous proliferation of the chemical marketplace, resulting in a vast array of new chemicals. To date, over 350,000 chemical substances have been registered for use in the global market.[11] While society benefits from the application of these new chemicals in everyday and economic contexts, the true economic and societal costs are generally externalized, as outlined in Box 1.[1,2] The responsibility of identifying hazardous chemicals typically falls on government agencies, academic scientists, and civil societal organizations, often occurring years or decades after their first use, while the responsibility of water treatment and remediation falls to water utilities, paid for by consumers.[12]

While a remarkable shift has taken place in the 20th century, from the failed "solving pollution by dilution" to building and developing centralized water-treatment facilities, identifying and understanding chemical pollutants and subsequent development of treatment technologies have been on a sporadic path, often only after environmental incidences have occurred.

Box 1. True economic costs of chemical water pollution

History shows the externalized economic and societal costs due to the widespread presence of hazardous chemicals in the environment are substantial and often not recognized until long after the compounds are in use. Put another way, these are the costs that the integrated interdisciplinary framework proposed by this perspective seeks to avoid or minimize.

Treatment and remediation costs: these are the expenses of treating drinking water to remove toxic chemicals or the costs of avoiding contamination or remediation of the water source. For example, to treat elevated levels of PFASs in the Cape Fear River Watershed, two drinking water-treatment plants were upgraded with granulated activated carbon and reverse osmosis systems at an expense of over $200 million. These costs have been borne by the utility ratepayers and taxpayers pending legal cases against the producers of the substances. In addition, the United States Department of Defense (DoD) estimated costs of cleanups related to their firefighting operations at greater than $2 billion. As of 2021, the DoD already spent $1.02 billion.[13,14]

Health effects: these are economic damages due to adverse health effects, typically calculated as a function of increased mortality or decreased quality of life. For instance, a recent study identified PFAS-attributable disease costs in the United States of $5.5 to $63 billion annually.[15] As our comprehension of the impact of chemical pollutants on human health advances, coupled with an enhanced understanding of the prevalence of pollutants in the environment, it is certain that more health costs will be uncovered in the future. The United States EPA uses "value of a statistical life" to assign monetary damages to changes in risks to human health because of exposure to chemicals. These values can be used to quantify the willingness to pay for people to reduce their risk of death.

Property price effects: many of the damages of chemicals go beyond health or treatment costs. Home prices can be affected by the discovery of hazardous substances near homes. The hedonic price techniques require comparing home sales near and far from contaminated sites and/or before and after contamination or cleanup. For example, studies show that leaking underground storage tanks are associated with a reduction in nearby home values by as much as 10%, but the effect was limited to those more severe and well-known cases.[16]

Ecosystems and biodiversity loss: The Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) has identified pollution as one of the five major drivers of global biodiversity loss.[17] The cost of ecosystem function and biodiversity loss can be quantified through economic methods of non-market valuation such as state preference surveys or revealed preference studies of behavior, which are related to people's values for those changes, to estimate damages to society for the presence of chemicals and their effect on the environment.

Social justice: oftentimes, water quality is more degraded in areas of lower income,[18] e.g., lead in public water systems[19] or acidic mine drainage in rural communities downstream of abandoned mines,[20] aggravating societal inequity in terms of quality of life, manifested across the issues of human health,[21] property value, and ecosystem health, thereby negatively impacting society as a whole.

For example, eutrophication in lakes, rivers, and coastal waters was widely recognized as a water pollution problem in the mid-20th century,[22] motivating technological developments to remove excessive nutrients (e.g., nitrogen, phosphorus) from wastewater in the 1950s–1970s.[23] Meanwhile, the rapid expansion of industrialization and commercial agriculture introduced a broad variety of emerging water pollutants that largely cannot be removed by traditional treatment technologies designed for nutrients (e.g., pesticides, industrial chemicals, disinfection by-products, surfactants, pharmaceuticals, X-ray contrast media, water-soluble polymers, microplastics).[24] Due to their often lower-level presence (pg/L to μg/L) compared with nutrient and metal contamination (μg/L to mg/L), such water pollutants have been commonly termed "micropollutants." Their presence, even at trace levels, can disturb aquatic ecosystems and has been linked to various adverse health effects.[24] With advances in analytical instrumentation, micropollutants only gained widespread appreciation as a major water pollution issue in the early 2000s, prompting ongoing development of more advanced water-treatment technologies.[25]

Progress is being made in identifying the presence of micropollutants in environmental media and developing technologies to treat them, but relying solely on a reactive approach is far from enough.[26,27] Despite advancements in analytical capabilities, current methods can only measure a fraction of the over 350,000 registered chemicals and mixtures present in the global market.[11,28] Additionally, the existence of numerous transformation products further compounds the challenge of detecting all water pollutants within a reasonable time frame. Our knowledge about micropollutants is also still evolving, both in terms of their hazards and treatability. Recent advancements in ecotoxicology, ecology, and analytical chemistry will likely further expand the scope of the number of compounds of interest, as well as the types of interrelated environmental effects that remain poorly understood.[29,30] This has been recently demonstrated by the introduction of new hazard classifications of persistent, mobile, and toxic (PMT) and very persistent and very mobile (vPvM) substances in the European Union (EU), which are technically challenging and expensive to remove from water.[31,32] It can be expected that in the future, we will identify further micropollutants that may require novel chemical-specific treatment technologies (e.g., chlorinated paraffins[33] and water-soluble polymers[34]). Relying solely on end-of-pipe water-treatment solutions is expensive and only effective to a certain extent. For example, Switzerland is investing $1.2 billion to upgrade 100 large wastewater-treatment plants by 2040 to eliminate over 80% of micropollutants from wastewater.[35]

Therefore, it is essential to adopt complementary proactive approaches to screening existing chemicals in global markets, identifying potential micropollutants, and informing

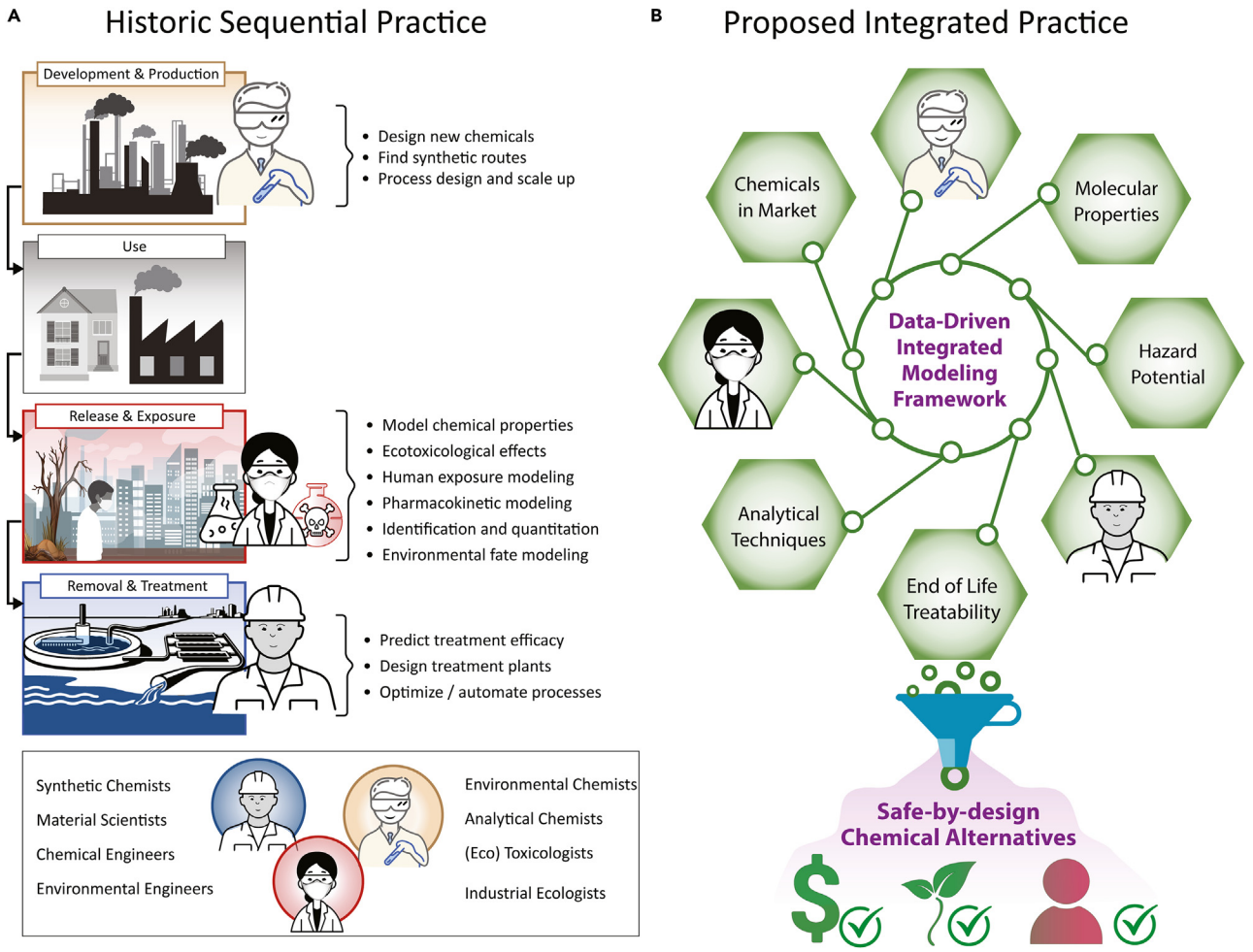

Figure 2. Comparison between historic and proposed frameworks

Historic sequential siloed practice following new chemical development (A) versus the proposed integrated practice (B) for modeling chemical behavior in different domains, including chemical development and production (yellow), release and exposure (red), and removal and treatment (blue), involving experts with varying backgrounds.

water-treatment technologies. Additionally, proactive design of novel chemicals that are benign and easily treatable is crucial. To achieve this, we call on researchers and practitioners in all relevant domains along the chemical life cycle (i.e., raw material acquisition, chemical production, use, end of life including water treatment, and environmental research)—that are currently siloed—to work together by using a novel data-driven integrated framework (Figure 2). This perspective starts with an analysis of how the individual domains are currently siloed. Learning from historical shortcomings, we elaborate on how a data-driven integrated framework can be built upon existing efforts in individual domains and how each domain may benefit from the proposed integrated framework. Finally, we provide recommendations on next steps to establish the proposed integrated framework.

EVERYTHING IS SILOED

Historically, the inability to foresee the impacts of emerging water pollutants has led to unwanted outcomes, inevitably following the same four-step sequence. Step 1: environmental and human exposure awareness, step 2: source identification, step 3: mitigation/treatment process development, and step 4: development and use of (often regrettable) replacements. Generally, each step is undertaken by experts from different domains of the chemical life cycle. This sequence has been repeated throughout history, as demonstrated by dry-cleaning solvents,[36] bisphenol A,[37] and, most recently, PFASs.[28] In all these examples, the same outcomes resulted: (1) continued environmental and human exposure while developing more effective treatment processes and (2) limited incorporation of lessons learned toward developing safer alternatives.

Taking perfluorooctanoic acid (PFOA, or C8) as an illustrative case, this highly stable compound possessing exceptional surfactant characteristics and other related properties has found widespread uses, including its utilization as a processing aid in the manufacturing of specific fluoropolymers.[38] First industrialized in the 1950s, due to limited and proprietary information, PFOA remained unnoticed in the public domain for several decades. Only in the late 1990s and early 2000s were PFOA and some other PFASs discovered by environmental scientists to be

ubiquitous in the global environment due to the advancements of analytical technologies (step 1).[39] Such awakening then triggered a wider community of environmental chemists and (eco)toxicologists (step 2) to spend nearly a decade concluding that PFOA is highly persistent in the environment, bioaccumulative, and toxic[40] while elucidating sources in the environment.[38,41,42] Informed by scientific advances, policy was developed to address the releases of PFOA. In 2019, PFOA was finally included in the Stockholm Convention on Persistence Organic Pollutants for elimination.[43] However, environmental engineers are still working on identifying effective and efficient methods to remove and destroy PFOA from contaminated sites and waters, which has proven to be very challenging (step 3).[44] Meanwhile, fluoropolymer manufacturers have replaced PFOA with alternative PFASs, including hexafluoropropylene oxide dimer acid (HFPO-DA, or GenX), with one major criterion being that the alternative "fulfills all requirements for the polymerization with minor process adjustments."[45] However, HFPO-DA has also shown to be toxic with similar adverse effects[46,47] while also being more difficult to remove from water compared with PFOA (step 4).[48] It should be highlighted that we described the situation of only one PFAS and its replacement (i.e., PFOA and HFPO-DA). The magnitude of the problem expands significantly when considering the hundreds of thousands of chemicals with different structures and properties being used or created by the chemical industry.[28] Unfortunately, most of these chemicals are poorly understood by external groups due to the lack of disclosure requirements of manufacturers.

Had the developers of PFOA and HFPO-DA and their use in fluoropolymer production looked past their silo at early stages, better design of chemicals and processes could have been achieved through joint effort with environmental chemists, (eco)toxicologists, and environmental engineers. This collaboration might have prevented the widespread use and release of PFOA and HFPO-DA and instead facilitated the development of safer alternatives. However, this is difficult to achieve in practice under the working framework so far.

Typically, chemical and process developers, environmental scientists, environmental engineers, and others are driven by different motivations, reacting myopically to their own immediate objectives with limited regard to other domains, effectively placing them and their outcomes in "silos" (Figure 2A). So here we find ourselves today, in stasis, with the chemical-discovery scientists developing new chemicals at unprecedented rates, while the identities of most chemicals put on the market remain unknown to other domains.[11] In turn, vast public resources are expended by environmental chemists to discover the presence of these "new" chemicals and their transformation products in the environment, often after years, or even decades, of widespread use. Following discoveries by environmental chemists, (eco)toxicologists assess the harmfulness of these "new" chemicals, while environmental engineers investigate the efficacy of their systems to remove or degrade the chemicals. While each of these experts continue to develop drastic technological advances in addressing problems in their respective domains, in the larger scheme, this reactive system fosters an ongoing pattern of the solution of one problem engendering challenges in allied domains along the chemical life cycle.

Building blocks to bridge diverse domains already exist

To break the four-step reactive sequence, we propose a proactive approach wherein efforts proceed in parallel among these domains, facilitated by a data-driven integrated framework. In particular, cheminformatics—or the use of chemistry data for information exchange and modeling—can reconcile the present disconnected data flows between siloed domains, providing alternative problem-solving approaches while generating new perspectives and solutions (Figure 2B).

Many models are already in use at various stages of the chemical life cycle, including design, synthesis, processing, environmental fate, (eco)toxicity, end-of-life treatment, and remediation (see Table S1 for examples). It should be noted that Table S1 and its descriptions are not exhaustive but are meant to show and inform scientists and practitioners about opportunities and potential collaboration that can be initiated across domains throughout the chemical life cycle.

Of these models, reactivity and property prediction (e.g., pesticidal activity, surfactant functionality, water resistance) are used by synthetic chemists in early stages of molecular design to inform selection of chemical candidates for further testing and development.

Generally, models for estimating properties related to environmental fate and (eco)toxicity of a chemical become the focus in later life cycle stages. For example, data mining, (quantitative) structural-activity relationships, and novel machine-/deep-learning approaches have found wide application in aquatic-toxicity databases, chemometric software, and other web-based applications.[2] Various modeling techniques and tools have also been developed to study the physicochemical properties, sources, fate, and transport of chemicals in the environment, as well as those associated wildlife exposure.[49,50]

For the latter remediation/treatment processes, environmental engineers have developed models to predict the efficacy of current processes for certain chemicals to inform designing future treatment processes.[51,52] Also, materials scientists, environmental chemists, and environmental engineers have successfully used cheminformatic approaches to develop novel treatment techniques and materials,[53] including novel sorption substrates for potential use in wastewater-treatment plants.[6]

Despite different research interests, it is interesting to highlight a common modeling approach that links specific endpoints of interest to chemical structures. Multiple endpoints are modeled by experts across different domains but use different names for similar concepts. For example, scientists in the pharmaceutical industry may refer to "log *P*" and pharmacokinetics, while environmental chemists and (eco)toxicologists may use terms like "log K_{OW}" (octanol-water partition coefficient) and toxicokinetics, respectively. The recent rise of advanced artificial intelligence (AI) and big-data analysis such as machine learning suggests that the seeds of uniting different domains already exist, including data exchange and joint modeling of similar (but at times differently named) endpoints and provision of data input to one another.[54]

Cheminformatics can benefit all domains

Applying the proposed integrated data-driven framework across domains not only supports scientists and practitioners to enhance their existing work but also creates synergistic

opportunities for novel work that bring co-benefits across domains.

DEVELOPMENT AND PRODUCTION DOMAIN

This domain supports the design of new chemicals and materials with reduced hazards and/or improved end-of-life treatability. Key experts within the domain include synthetic chemists, materials scientists, chemical engineers, and process engineers.

Producers of new chemicals operate under market, economic, and regulatory constraints. Cheminformatics can inform the strategic directions and feasibility of new innovations by simultaneously tuning the design to meet goals of functionality and safe and sustainable by design.[55] To achieve safe by design, considerations of hazardous properties and end-of-life treatability need to be incorporated early on in the molecular design. Notably, empirical (eco)toxicological, environmental fate, and treatability studies can be costly. Therefore, early *in silico* screening and informed molecular design of the chemicals that may later move on to more detailed analysis saves valuable research efforts and resources.

Hence, it is crucial to integrate modeling approaches and tools developed by environmental chemists, (eco)toxicologists, and environmental engineers into current molecular design and retrosynthesis prediction tools that focus primarily on the functionality.[56] For example, structural-feature-based models[57] may be used to avoid or select specific chemical substructures that may cause undesired biological activities. This approach has been done in the drug-design field at a limited scale[58] but needs to be expanded to more general use for new chemical development and to include chemical substructures that may lead to other hazardous properties such as persistence, bioaccumulation, and mobility.[59,60] Another clear example can be found in the pesticides industry, where numerous pesticides exhibit chirality. For ease of manufacture, pesticides are commonly manufactured as a racemic mix, despite only one enantiomer being primarily bioactive and the other enantiomer(s) being excess ballast or not delivering the intended outcome.[61] Cheminformatics could inform whether these unintended enantiomers have unintentional bioactive properties in the environment (e.g., acting on non-target species), consequently requiring their removal before market introduction. Furthermore, the integration of such models into retrosynthesis prediction tools can help to select optimal synthesis routes for the target compounds that have minimized hazardous by-products.

Assessing the environmental impacts of pesticide transformation products is a common regulatory necessity in many parts of the world[62] and a rarity for most other chemical classes. To determine transformation products, tests under relevant environmental conditions (e.g., combinations of temperature, microbial communities, and UV radiation) are conducted, which is time consuming and resource intensive. Thus, it is not economically favorable for many industries to understand the environmental transformation products of their chemicals. The absence of these rigorous considerations has been shown to cause unexpected harmful effects for extended periods. For example, a transformation product (6PPD-quinone) of a common antioxidant in tires, 6PPD, has only recently been identified for causing severe acute death of wild fish populations. It required skillful and specialized environmental scientists nearly a decade of detailed forensic analysis to identify this impact.[63,64] Thus, it is highly desirable for early understanding and communication of transformation products by the development and production domain.

Although identifying potentially hazardous transformation products of industrial chemicals is challenging, recent advancements in analytical technologies coupled with data-mining techniques (e.g., suspect screening and non-targeted analysis) in environmental forensics have paved the way to get robust and relatively fast protocols to accomplish these tasks (see below). Further model developments are needed to facilitate enhanced chemical structure-based prediction of degradation pathways, transformation product identification, and subsequent assessment of transformation products. Considerable efforts are currently underway to predict reaction and synthetic pathways, which can also be extended to predict transformation products, offering valuable insights in this regard.[65–67]

RELEASE AND DETECTION DOMAIN

This domain supports understanding the release, exposure, and effects of existing chemicals. Key experts within the domain include environmental chemists, analytical chemists, (eco)toxicologists, and industrial ecologists.

In parallel to designing new safe-and-sustainable-by-design chemicals and materials, existing chemicals in use, as well as their transformation products, still need to be screened, prioritized, and assessed. To achieve this in an efficient and timely manner, key experts can benefit from data and modeling approaches generated by experts in adjacent (sub)domains.

Understanding chemicals in use

Environmental chemists and (eco)toxicologists may learn from practitioners in the development and production domain regarding the traits of chemicals on the global market (e.g., chemical series, lead compounds, bioactive moieties, and scaffolds), as well as possible impurities and by-products in their respective formulations. Such knowledge can help, for example, in defining the scope of existing chemicals for *in silico* hazard screening.[68] Among others, this expert knowledge could be leveraged to develop novel methods that help elucidate the chemical identities of environmental unknowns, particularly for the most challenging, long-overlooked ones such as UVCBs (unknown or variable composition, complex reaction products, or biological materials) and polymers.[69,70] By transitioning from current ambiguous substance names and descriptions to more explicit structural translation, these novel methods facilitate a more comprehensive understanding of these substances, paving ways for developing various models mentioned above and below for them.[71–73]

Supporting substance identification using non-targeted analytical techniques

The increasing utilization of high-resolution mass spectrometry for analyzing environmental samples has significantly expanded the role of cheminformatics for non-targeted investigations.[74,75] Previously, investigations into environmental transformation products with targeted methods required a labor-intensive and iterative process. Starting with known parent compounds, chemists would first hypothesize possible transformation

products,[76] followed by evaluating the hypothetical products with regard to likely collision-induced molecular fragments and thus developing a method to detect the suspected fragments. If no compounds were detected, chemists had to repeat the process of hypothesizing, evaluating, and detecting for other potential transformation schemes.[77,78] Such a trial-and-error approach often required months of effort to perhaps identify one or two transformation products. With the advent of non-targeted instrumental capabilities, cheminformatics routinely is invoked as the first step in searching the analytical data for molecular features recorded in multiple online databases (Table S1). This enables tentative identification of previously reported compounds, including those by synthetic chemists, via database matching. This approach, known as "suspect screening," has been supported by open data sources on chemical identities and existing measurements, such as the NORMAN Suspect List Exchange.[79] Following such online searches, online transformation models[76,80] developed by environmental chemists can also be invoked to propose hypothetical transformation schemes. These scheme can be used to interrogate the non-target analytical data and tentatively identify whether the proposed transformation products are present.[81]

Studying environmental fate and transport

Currently, the parameterization of chemicals is a major bottleneck when developing robust predictive models for contaminant fate in soil and sediment, particularly for ionizable and ionic compounds (e.g., most PMT and vPvM compounds).[82,83] This challenge arises due to several factors. First, these media are highly complex, with different mechanisms playing major roles at different layers. For example, with increasing depth through the vadose zone and into aquifers, organic matter content can drop orders of magnitude, and adsorptive mineral(oid)s such as hydrous ferric and aluminum oxides play an increasing role.[28,82] Second, existing approaches were developed for neutral organic compounds and are unable to describe the complex sorption behavior of ionic or ionizable compounds.[72] To overcome these limitations, collaborations among experts from various domains are essential. For instance, by incorporating molecular descriptors employed in the development domain, new predictive models can be parameterized, thereby leveraging the expertise of researchers specializing in molecular descriptors. Such multidisciplinary collaborations hold the potential for unraveling the capabilities of cheminformatics in elucidating and predicting interactions between chemicals, considering spatially variable conditions and phases at a mechanistic level.[84] Additionally, future models can be improved and validated by profiting off the increased volume of data resulting from advancements in targeted and non-targeted analytics. Such advancements facilitate tracking of novel contaminants, identifying transformation products, and determining their persistence in the environment. By leveraging these extensive datasets, predictive models can be refined and improved, bolstering the ability to comprehend and forecast the behavior of diverse chemicals in complex environmental systems.[83]

Modeling biological behavior and effects

Understanding the biological behavior and effects of chemicals is challenged by the complexity of biology. In the drug-discovery field, physiologically based pharmacokinetic (PBPK) modeling has been successfully used to understand absorption, distribution, metabolism, and excretion of chemicals in humans and animals, with large amounts of empirical data and different modeling approaches established for many chemicals.[85,86] Notably, most of these data and modeling approaches are proprietary information of pharmaceutical companies. Transitioning toward data accessibility and standardization in PBPK modeling and reporting is crucial to drive interdisciplinary collaboration across the chemical life cycle.[87] By making these resources more widely available and consistent, synergistic benefits can be achieved, enabling researchers to capitalize on the accumulated knowledge, enhance the understanding of bioaccumulation, and facilitate (eco)toxicological modeling. Embracing data accessibility and homogeneity in PBPK modeling will not only promote scientific collaboration but will also accelerate advancements in the understanding of chemical behavior within biological systems.

REMOVAL AND TREATMENT DOMAIN

This domain supports the development of new removal techniques and treatment plants. Key experts within the domain include environmental engineers, civil engineers, synthetic chemists, materials scientists, and analytical chemists.

Water treatment aims to maximize efficiency and minimize cost via engineered processes. However, the increasing presence of diverse anthropogenic chemicals in the environment, coupled with the adverse impacts of climate change and land use alterations on the hydrological cycle (e.g., droughts, floods), has exacerbated challenges faced in water treatment.[88,89] Thus, development of novel treatment technologies is imperative.[90] Likewise, treatment facilities are being reimagined as water-resource recovery facilities where water, nutrients, and energy can be recovered and monetized, contributing toward advancements in the water-energy-food security nexus.[91] To achieve the highest efficiency at the lowest cost, computer-aided techniques are essential for water treatment. Such techniques can be leveraged for technology selection or treatment combination (Table S1).[92] However, this strategy remains underexplored. Here, we provide two examples to illustrate the potential of using cheminformatics to drive discovery.

Predicting and managing water-treatment operations

Water-treatment facilities are designed and engineered to meet well-defined specifications (i.e., water-quality parameters). The efficiency in treatment operations depends largely on the influent water quality, requiring monitoring, detection, and removal of various chemicals. However, these tasks require significant resources, such as time, energy, and financial investment, positing a burden on public utilities.[93] Combining the cheminformatics advanced modeling capabilities, civil and environmental engineers can work with environmental chemists and (eco)toxicologists to optimize treatment operations (e.g., adsorbent and chemicals use, energy consumption), reducing the financial burden on an already fiscally limited budget.[94] Establishing better data practices, as explained in the following section, will enhance the availability of knowledge and data for civil and environmental engineers to be prepared for emerging chemicals of concern.[95] For example, predicting chemical transformations in the environment will equip civil and environmental engineers to design and

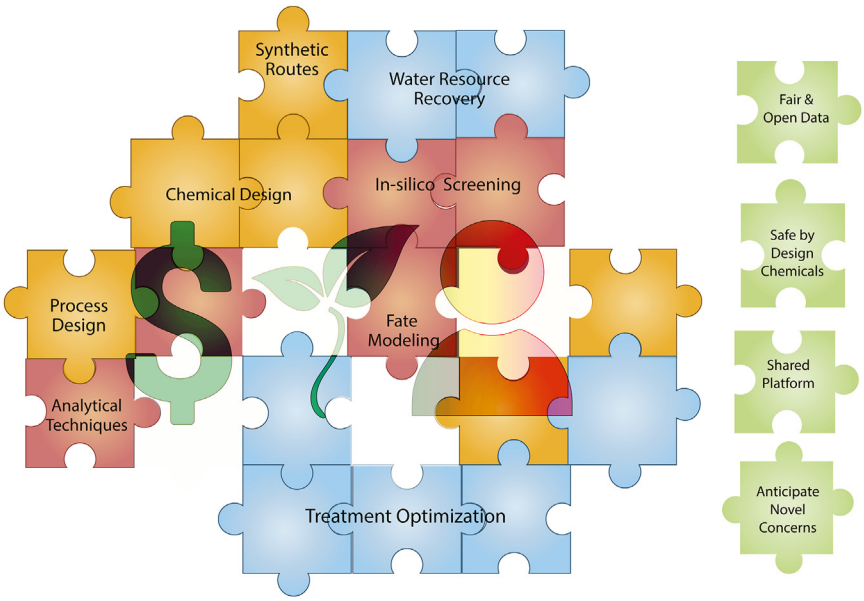

Figure 3. Recommended actions
Concerted actions to operationalize the integrated data-driven framework to support safer, sustainable, and economically feasible decisions.

manage water-treatment facilities more effectively in the future. This includes considerations of physical adsorbents, membrane separations, and chemical treatment processes, as well as capital investments, operational strategies, and maintenance protocols.[96–98] Likewise, as environmental chemists and (eco)toxicologists develop enhanced fate and transformations models, civil and environmental engineers can better predict outcomes from severe weather events (e.g., drought, flooding).[93] By embracing interdisciplinary collaboration and leveraging models and data provided by cheminformatics, water treatment can continuously improve its operations and resilience.

Next-generation water-resource recovery facilities

Energy generators, such as coal plants, have undergone retrofits to house miniature water-treatment facilities that not only ensure compliance with water-quality standards but also capture carbon, providing an additional revenue stream.[99] Developed by coal power plants, this model provides a framework that can be utilized by materials scientists and engineers, chemical engineers, and civil/environmental engineers. Dilute aqueous streams, processed for water quality, provide a raw resource that can be mined for valuable chemicals such as ammonia,[100] phosphate,[101] carbon dioxide,[102] and precious metals,[103] as well as energy.[104] By enhancing operational abilities of treatment facilities, the integration of chemical processing and energy generation becomes feasible, providing new revenue streams. Meanwhile, better understanding and preparation for periodic weather events (e.g., algae blooms) allows treatment facilities to respond appropriately, ensuring continuous water quality while taking advantage of a variable raw feedstock (i.e., water influent).

CONCERTED ACTIONS TO OPERATIONALIZE THE PROPOSED FRAMEWORK

To operationalize the proposed framework, the following subsections outline several concerted actions warranted by the wider scientific community (Figure 3). Right now, the ability to predict future hazards, exposure, and risks of chemicals is limited. Therefore, while implementing cross-domain cheminformatics can be a game changer, the lessons of the past should temper expectations of it being a replacement for the long-term monitoring and continuous discovery of new facets regarding fate and effects of chemicals in the environment.[105]

Applying open and FAIR data practices

The foundation of our proposed integrated framework starts with data, much like other modeling approaches. Data on chemicals and their properties will effectively serve as the common currency among experts seeking to collaborate on modeling efforts across the different domains and stages of the chemical life cycle, with data availability, quality, and suitability as key factors when selecting among modeling techniques. However, the current large discrepancies in format and accessibility of the above-mentioned data and models are a major obstacle.

Therefore, all stakeholders, particularly those generating, curating, and maintaining data, should ensure that these data are findable, accessible, interoperable, and reusable (FAIR).[106,107] This can start with using established open chemical identifiers and formats, such as InChI (International Chemical Identifier), which provide harmonization and uniqueness, as opposed to closed and non-unique formats like CASRN (Chemical Abstracts Service Registry Number). In addition, forums and projects, similarly to the WorldFAIR project,[108] need to be established to convene experts across domains to discuss and standardize terminologies, data format, and other relevant topics ensuring interoperability. To maximize the (re)usability, transparency, and utility of data and models, they should also be open, that is, accessible to and reusable for research and assessment purposes (instead of commercial, profit-making enterprises). Specifically, input data, model parameters, and software tools should be disclosed and made available to ensure maximum reproducibility and foster bridging across the current silos so that scientists working in one domain can easily use the data/models generated by those from another.

To achieve this goal, it is evident that a cross-disciplinary exchange of harmonized data and process-level understanding that works toward an open-architecture transdisciplinary chemical design and assessment is needed. Encouragement by regulators, funding agencies, and academia of this open-architecture transdisciplinary approach could foster the transition to a common framework. For example, these next-generation aspirations are among the specific objectives of a major new research program of the United States Environmental Protection Agency (EPA).[109] Initiated in 2022, this major research collaboration

aims to advance the interoperability of cheminformatics and modeling across disciplines to bring unprecedented capabilities to the United States EPA's new chemical assessment and registration process. With a similar motivation, the EU's Action Plan "Toward Zero Pollution for Air, Water and Soil" acknowledges that "we need a better integrated overview of pollution for public and private actors to tackle connected pollution issues across space and time and address their interplay with other environmental, social and economic considerations as effectively as possible in their policy, investment and purchase decisions" and embraces safe-and-sustainable-by-design principles for the development of new chemicals. These initiatives need to be upscaled.[110]

Develop common knowledge bases/platforms

Recently, Google and Apple have teamed up with the non-governmental organization (NGO) ChemForward to take a proactive approach and address the PFAS issue in the electronics industry.[111] In brief, ChemForward has adopted an integrative and inclusive strategy that moves beyond the traditional siloed assessment of new compounds as described in Figure 1A. Specifically, they accomplish this by establishing a centralized peer-reviewed repository with one single hazard assessment per chemical, harmonizing findings and parameters from different sources and methodologies. Each assessment contains an evaluation and classification of human health hazards, physical-chemical properties, and environmental fate and toxicity endpoints based on criteria that are communicated transparently. It remains, however, too early to see how successful this initiative will be in developing high-performing and safer PFAS-free alternatives and, beyond that, actually establishing such alternatives as an industry standard, which often can be an additional bottleneck. Nevertheless, to develop cross-disciplinary models, the availability of trusted and harmonized data repositories is essential, making the approach by ChemForward a meaningful step toward creating unified transdisciplinary models for chemical assessment and design.

Invoke transdisciplinary input as early as possible in the life cycle

The proposed integrated framework can be applied to design and develop new chemicals to prevent future water pollution while supporting the assessment and sound management of existing chemicals in use to minimize or eliminate water pollution. Early application of the framework in the chemical development phase would be most effective, minimizing long-term harm, including regrettable substitution, externalized costs to health and the environment (see Box 1), and preventing "lock in."[112] In decades past, informed contributions from downstream domains would have required considerable time and financial investment to conduct a wide array of experiments on factors such as environmental persistence, bioaccumulation, mobility, toxicity to model species, and efficacy of existing treatment technologies. The risk of a market niche being lost to a competitor during this protracted testing adds more opportunity costs. However, the advanced state of today's cheminformatics and modeling systems, as proposed by this perspective, renders this valuable information for preselection of candidate structures at low cost but potentially yields immense positive value in the long run for industry, the environment, and society as a whole.

Recognizing novel "properties" of concern in a vigilant manner

To develop future-proof modeling frameworks, we need to make sure that our harmonized, accessible databases and models are constructed in an architecture that allows for additional factors to be included when new necessities arise. For example, over the last few decades, a chemical's ability to disrupt the endocrine system of living organisms has gained increasing attention. This mode of toxic action is generally more difficult to determine experimentally, as the effects are subtler and less acute compared with other toxic effects commonly determined in (eco)toxicological screenings. In 2009, the EU included endocrine disruption as a knockout criteria for the market admission of pesticides.[113] This caused a non-approval of admission for several pesticides, including mancozeb and tolpyralate.[114,115] The reactive action in this case shows that the scientific, regulatory, and societal assessment on what constitutes a risk or a chemical property of concern will evolve over time. Thus, it is crucial to account for this evolution when developing unified transdisciplinary modeling approaches within the open and FAIR architecture. Doing so will ensure that novel frameworks and models remain adaptable and capable of accommodating emerging concerns and knowledge in a timely manner.[95,116]

Finally, with the ability to screen chemicals more holistically and efficiently for their potential impact across their life cycle, society will be moving in the direction of safe and sustainable by design and will avoid the damages that are typically realized only many years after a chemical's widespread use. Striving for enhanced transdisciplinary collaborations, common data standards and interoperable information systems are worth the concerted effort by experts across domains. The avoided costs to society from initial cleanups to long-term health impacts and ecological damages far outweigh the frictions in moving our chemical-information systems toward a more forward-leaning and proactive framework.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.oneear.2023.07.001.

ACKNOWLEDGMENTS

Z.W. gratefully acknowledges funding from the EU under the Horizon 2020 Research and Innovation Program (grant agreement number 101036756). A.L. acknowledges support from the Luxembourg National Research Fund (FNR) for Project A18/BM/12341006. This document has been subjected to the United States EPA's review and has been approved for publication. The views expressed in this article are those of the authors and do not necessarily represent the views or policies of the agency. Any mention of trade names, products, or services does not imply an endorsement by the agency. The agency does not endorse any commercial products, services, or enterprises.

AUTHOR CONTRIBUTIONS

Conceptualization, M.A. and Z.W.; writing – original draft, M.A., G.S., M.J.B., J.W.W., A.L., N.H.M., and Z.W.; writing – review & editing, M.A., G.S., M.J.B., J.W.W., A.L., N.H.M., and Z.W.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Backhaus, T., Snape, J., and Lazorchak, J. (2012). The impact of chemical pollution on biodiversity and ecosystem services: the need for an improved understanding. Integrated Environ. Assess. Manag. *8*, 575–576.
2. Groh, K., Vom Berg, C., Schirmer, K., and Tlili, A. (2022). Anthropogenic Chemicals As Underestimated Drivers of Biodiversity Loss: Scientific and Societal Implications. Environ. Sci. Technol. *56*, 707–710.
3. Andrews, D.Q., and Naidenko, O.V. (2020). Population-wide exposure to per-and polyfluoroalkyl substances from drinking water in the United States. Environ. Sci. Technol. Lett. *7*, 931–936.
4. Wilkinson, J.L., Boxall, A.B.A., Kolpin, D.W., Leung, K.M.Y., Lai, R.W.S., Galbán-Malagón, C., Adell, A.D., Mondon, J., Metian, M., Marchant, R.A., et al. (2022). Pharmaceutical pollution of the world's rivers. Proc. Natl. Acad. Sci. USA *119*, e2113947119.
5. Baste, I.A., Watson, R.T., Brauman, K.I., Samper, C., and Walzer, C. (2021). Making Peace with Nature: A Scientific Blueprint to Tackle the Climate, Biodiversity and Pollution Emergencies. https://www.unep.org/resources/making-peace-nature.
6. Podgorski, J., and Berg, M. (2020). Global threat of arsenic in groundwater. Science *368*, 845–850.
7. Wang, Z., and Hellweg, S. (2021). First steps toward sustainable circular uses of chemicals: advancing the assessment and management paradigm. ACS Sustain. Chem. Eng. *9*, 6939–6951.
8. CAS (2023). CAS Registry Number. https://www.cas.org/support.
9. Fuller, R., Landrigan, P.J., Balakrishnan, K., Bathan, G., Bose-O'Reilly, S., Brauer, M., Caravanos, J., Chiles, T., Cohen, A., Corra, L., et al. (2022). Pollution and health: a progress update. Lancet Planet. Health *6*. e535-e547.
10. Alabaster, G., Johnston, R., Thevenon, F., and Shantz, A. (2021). Progress on Wastewater Treatment Global Status and Acceleration Needs for SDG Indicator 6.3. 1. https://www.unwater.org/publications/progress-wastewater-treatment-2021-update.
11. Wang, Z., Walker, G.W., Muir, D.C.G., and Nagatani-Yoshida, K. (2020). Toward a global understanding of chemical pollution: a first comprehensive analysis of national and regional chemical inventories. Environ. Sci. Technol. *54*, 2575–2584.
12. Cordner, A., Goldenman, G., Birnbaum, L.S., Brown, P., Miller, M.F., Mueller, R., Patton, S., Salvatore, D.H., and Trasande, L. (2021). The true cost of PFAS and the benefits of acting now. Environ. Sci. Technol. *55*, 9630–9633.
13. US DoD (2021). Officials Describe DOD's Efforts to Mitigate Impacts of PFAS Chemicals. https://www.defense.gov/News/News-Stories/Article/Article/2869455/officials-describe-dods-efforts-to-mitigate-impacts-of-pfas-chemicals/.
14. US DoD (2022). Report on Per- and Polyfluoroalkyl Substances Active Sites Cleanup Costs. https://media.defense.gov/2022/Jul/21/2003039883/-1/-1/1/REPORT-ON-PER-AND-POLYFLUOROALKYL-SUBSTANCES-ACTIVE-SITES-CLEANUP-COSTS-June-2022.PDF.
15. Obsekov, V., Kahn, L.G., and Trasande, L. (2023). Leveraging Systematic Reviews to Explore Disease Burden and Costs of Per-and Polyfluoroalkyl Substance Exposures in the United States. Expo. Health *15*, 373–394.
16. Zabel, J.E., and Guignet, D. (2012). A hedonic analysis of the impact of LUST sites on house prices. Resour. Energy Econ. *34*, 549–564.
17. Díaz, S.M., Settele, J., Brondízio, E., Ngo, H., Guèze, M., Agard, J., Arneth, A., Balvanera, P., Brauman, K., Butchart, S., and Chan, K. (2019). The Global Assessment Report on Biodiversity and Ecosystem Services | Summary for Policy Makers. https://ipbes.net/sites/default/files/inline/files/ipbes_global_assessment_report_summary_for_policymakers.pdf.
18. Daneshvar, F., Nejadhashemi, A.P., Zhang, Z., Herman, M.R., Shortridge, A., and Marquart-Pyatt, S. (2016). Evaluating stream health based environmental justice model performance at different spatial scales. J. Hydrol. *538*, 500–514.
19. Whitehead, L.S., and Buchanan, S.D. (2019). Childhood lead poisoning: a perpetual environmental justice issue? J. Publ. Health Manag. Pract. *25*. S115-S120.
20. Moeng, K. (2019). Community perceptions on the health risks of acid mine drainage: the environmental justice struggles of communities near mining fields. Environment. Environ. Dev. Sustain. *21*, 2619–2640.
21. Tariqi, A.Q., and Naughton, C.C. (2021). Water, health, and environmental justice in California: Geospatial analysis of nitrate contamination and thyroid cancer. Environ. Eng. Sci. *38*, 377–388.
22. Rodhe, W. (1969). Crystallization of eutrophication concepts in northern Europe. In Proceedings of Symposium on Eutrophication: Causes, Consequences, Correctives (National Academy of Sciences), pp. 50–64.
23. Lofrano, G., and Brown, J. (2010). Wastewater management through the ages: A history of mankind. Sci. Total Environ. *408*, 5254–5264.
24. Schwarzenbach, R.P., Escher, B.I., Fenner, K., Hofstetter, T.B., Johnson, C.A., Von Gunten, U., and Wehrli, B. (2006). The challenge of micropollutants in aquatic systems. Science *313*, 1072–1077.
25. Joss, A., Siegrist, H., and Ternes, T.A. (2008). Are we about to upgrade wastewater treatment for removing organic micropollutants? Water Sci. Technol. *57*, 251–255.
26. Wagner, W.E., and Gold, S.C. (2022). Legal obstacles to toxic chemical research. Science *375*, 138–141.
27. Orive, G., Lertxundi, U., Brodin, T., and Manning, P. (2022). Greening the pharmacy. Science *377*, 259–260.
28. Evich, M.G., Davis, M.J.B., McCord, J.P., Acrey, B., Awkerman, J.A., Knappe, D.R.U., Lindstrom, A.B., Speth, T.F., Tebes-Stevens, C., Strynar, M.J., et al. (2022). Per-and polyfluoroalkyl substances in the environment. Science *375*, eabg9065.
29. Sigmund, G., Ågerstrand, M., Antonelli, A., Backhaus, T., Brodin, T., Diamond, M.L., Erdelen, W.R., Evers, D.C., Hofmann, T., Hueffer, T., et al. (2023). Addressing chemical pollution in biodiversity research. Global Change Biol. *29*, 3240–3255.
30. Ateia, M., Buren, J.V., Barrett, W., Martin, T., and Back, G.G. (2023). Sunrise of PFAS Replacements: A Perspective on Fluorine-Free Foams. ACS Sustain. Chem. Eng. *11*, 7986–7996.
31. Hale, S.E., Arp, H.P.H., Schliebner, I., and Neumann, M. (2020). Persistent, mobile and toxic (PMT) and very persistent and very mobile (vPvM) substances pose an equivalent level of concern to persistent, bioaccumulative and toxic (PBT) and very persistent and very bioaccumulative (vPvB) substances under REACH. Environ. Sci. Eur. *32*, 155–215.
32. Hale, S.E., Neumann, M., Schliebner, I., Schulze, J., Averbeck, F.S., Castell-Exner, C., Collard, M., Drmač, D., Hartmann, J., Hofman-Caris, R., et al. (2022). Getting in control of persistent, mobile and toxic (PMT) and very persistent and very mobile (vPvM) substances to protect water resources: strategies from diverse perspectives. Environ. Sci. Eur. *34*, 22–24.
33. Chen, C., Chen, A., Zhan, F., Wania, F., Zhang, S., Li, L., and Liu, J. (2022). Global Historical Production, Use, In-Use Stocks, and Emissions of Short-Medium-and Long-Chain Chlorinated Paraffins. Environ. Sci. Technol. *56*, 7895–7904.
34. Arp, H.P.H., and Knutsen, H. (2020). Could we spare a moment of the spotlight for persistent, water-soluble polymers? Environ. Sci. Technol. *54*, 3–5.
35. Eggen, R.I.L., Hollender, J., Joss, A., Schärer, M., and Stamm, C. (2014). Reducing the discharge of micropollutants in the aquatic environment: the benefits of upgrading wastewater treatment plants. Environ. Sci. Technol. *48*, 7683–7689.
36. Lohman, J. (2002). A history of dry cleaners and sources of solvent releases from dry cleaning equipment. Environ. Forensics *3*, 35–58.
37. Huang, Y., Yin, H., Wang, J., Liu, Q., Wu, C., Wahlström, B., Neretin, L., and Wong, M.H. (2012). Bisphenol A (BPA) in China: a review of sources, environmental levels, and potential human health impacts. Gene X. *498*, 91–95.
38. Wang, Z., Cousins, I.T., Scheringer, M., Buck, R.C., and Hungerbühler, K. (2014). Global emission inventories for C4–C14 perfluoroalkyl carboxylic acid (PFCA) homologues from 1951 to 2030, Part I: production and emissions from quantifiable sources. Environ. Int. *70*, 62–75.
39. Fenton, S.E., Ducatman, A., Boobis, A., DeWitt, J.C., Lau, C., Ng, C., Smith, J.S., and Roberts, S.M. (2021). Per-and polyfluoroalkyl substance toxicity and human health review: Current state of knowledge and strategies for informing future research. Environ. Toxicol. Chem. *40*, 606–630.
40. Vierke, L., Staude, C., Biegel-Engler, A., Drost, W., and Schulte, C. (2012). Perfluorooctanoic acid (PFOA)—main concerns and regulatory developments in Europe from an environmental point of view. Environ. Sci. Eur. *24*. 16–11.
41. Prevedouros, K., Cousins, I.T., Buck, R.C., and Korzeniowski, S.H. (2006). Sources, fate and transport of perfluorocarboxylates. Environ. Sci. Technol. *40*, 32–44.
42. Wang, Z., Cousins, I.T., Scheringer, M., Buck, R.C., and Hungerbühler, K. (2014). Global emission inventories for C4–C14 perfluoroalkyl carboxylic acid (PFCA) homologues from 1951 to 2030, part II: The remaining pieces of the puzzle. Environ. Int. *69*, 166–176.

43. Stockholm Convention (2019). The New POPs under the Stockholm Convention - Website. http://chm.pops.int/Convention/ThePOPs/TheNewPOPs/tabid/2511/Default.aspx.
44. Wang, Z., DeWitt, J.C., Higgins, C.P., and Cousins, I.T. (2017). A never-ending story of per-and polyfluoroalkyl substances (PFASs)? Environ. Sci. Technol. *51*, 2508–2518.
45. Wang, Z., Cousins, I.T., Scheringer, M., and Hungerbühler, K. (2013). Fluorinated alternatives to long-chain perfluoroalkyl carboxylic acids (PFCAs), perfluoroalkane sulfonic acids (PFSAs) and their potential precursors. Environ. Int. *60*, 242–248.
46. US EPA (2021). Human Health Toxicity Assessments for GenX Chemicals - Website. https://www.epa.gov/chemical-research/human-health-toxicity-assessments-genx-chemicals.
47. The Intercept (2016). New TEFLON Toxin Causes Cancer in Lab Animals - Website. https://theintercept.com/2016/03/03/new-teflon-toxin-causes-cancer-in-lab-animals/.
48. Vakili, M., Bao, Y., Gholami, F., Gholami, Z., Deng, S., Wang, W., Kumar Awasthi, A., Rafatullah, M., Cagnetta, G., and Yu, G. (2021). Removal of HFPO-DA (GenX) from aqueous solutions: A mini-review. Chem. Eng. J. *424*, 130266.
49. Di Guardo, A., Gouin, T., MacLeod, M., and Scheringer, M. (2018). Environmental fate and exposure models: advances and challenges in 21 st century chemical risk assessment. Environ. Sci. Process. Impacts *20*, 58–71.
50. Awfa, D., Ateia, M., Mendoza, D., and Yoshimura, C. (2021). Application of Quantitative Structure–Property Relationship Predictive Models to Water Treatment: A Critical Review. ACS ES. T. Water *1*, 498–517.
51. Scharf, R.G., Johnston, R.W., Semmens, M.J., and Hozalski, R.M. (2010). Comparison of batch sorption tests, pilot studies, and modeling for estimating GAC bed life. Water Res. *44*, 769–780.
52. Jin, X., Tang, C.Y., Gu, Y., She, Q., and Qi, S. (2011). Boric acid permeation in forward osmosis membrane processes: modeling, experiments, and implications. Environ. Sci. Technol. *45*, 2323–2330.
53. Yang, J., Tao, L., He, J., McCutcheon, J.R., and Li, Y. (2022). Machine learning enables interpretable discovery of innovative polymers for gas separation membranes. Sci. Adv. *8*, eabn9545.
54. White, A.D. (2023). The future of chemistry is language. Nat. Rev. Chem, 1–2.
55. Pyzer-Knapp, E.O., Pitera, J.W., Staar, P.W.J., Takeda, S., Laino, T., Sanders, D.P., Sexton, J., Smith, J.R., and Curioni, A. (2022). Accelerating materials discovery using artificial intelligence, high performance computing and robotics. npj Comput. Mater. *8*, 84–89.
56. Elton, D.C., Boukouvalas, Z., Fuge, M.D., and Chung, P.W. (2019). Deep learning for molecular design—a review of the state of the art. Mol. Syst. Des. Eng. *4*, 828–849.
57. Houck, K.A., Patlewicz, G., Richard, A.M., Williams, A.J., Shobair, M.A., Smeltz, M., Clifton, M.S., Wetmore, B., Medvedev, A., and Makarov, S. (2021). Bioactivity profiling of per-and polyfluoroalkyl substances (PFAS) identifies potential toxicity pathways related to molecular structure. Toxicology *457*, 152789.
58. Limban, C., Nuță, D.C., Chiriță, C., Negreș, S., Arsene, A.L., Goumenou, M., Karakitsios, S.P., Tsatsakis, A.M., and Sarigiannis, D.A. (2018). The use of structural alerts to avoid the toxicity of pharmaceuticals. Toxicol Rep *5*, 943–953.
59. van Dijk, J., Flerlage, H., Beijer, S., Slootweg, J.C., and van Wezel, A.P. (2022). Safe and sustainable by design: A computer-based approach to redesign chemicals for reduced environmental hazards. Chemosphere *296*, 134050.
60. Suk, M., Lorenz, S., and Kümmerer, K. (2023). Identification of environmentally biodegradable scaffolds for the benign design of quinolones and related substances. Sustainable Chemistry and Pharmacy *31*, 100947.
61. Ulrich, E.M., TenBrook, P.L., McMillan, L.M., Wang, Q., and Lao, W. (2018). Enantiomer-specific measurements of current-use pesticides in aquatic systems. Environ. Toxicol. Chem. *37*, 99–106.
62. Latino, D.A.R.S., Wicker, J., Gütlein, M., Schmid, E., Kramer, S., and Fenner, K. (2017). Eawag-Soil in enviPath: a new resource for exploring regulatory pesticide soil biodegradation pathways and half-life data. Environ. Sci.*: Process. Impacts 19*, 449–464.
63. Tian, Z., Zhao, H., Peter, K.T., Gonzalez, M., Wetzel, J., Wu, C., Hu, X., Prat, J., Mudrock, E., Hettinger, R., et al. (2021). A ubiquitous tire rubber–derived chemical induces acute mortality in coho salmon. Science *371*, 185–189.
64. Tian, Z., Gonzalez, M., Rideout, C.A., Zhao, H.N., Hu, X., Wetzel, J., Mudrock, E., James, C.A., McIntyre, J.K., and Kolodziej, E.P. (2022). 6PPD-Quinone: Revised Toxicity Assessment and Quantification with a Commercial Standard. Environ. Sci. Technol. Lett. *9*, 140–146.
65. Simm, G.N., Vaucher, A.C., and Reiher, M. (2019). Exploration of reaction pathways and chemical transformation networks. J. Phys. Chem. A *123*, 385–399.
66. Engkvist, O., Norrby, P.-O., Selmi, N., Lam, Y.-h., Peng, Z., Sherer, E.C., Amberg, W., Erhard, T., and Smyth, L.A. (2018). Computational prediction of chemical reactions: current status and outlook. Drug Discov. Today *23*, 1203–1218.
67. Cuperlovic-Culf, M. (2018). Machine learning methods for analysis of metabolic data and metabolic pathway modeling. Metabolites *8*, 4.
68. Clark, A.M., Gedeck, P., Cheung, P.P., and Bunin, B.A. (2021). Using Machine Learning to Parse Chemical Mixture Descriptions. ACS Omega *6*, 22400–22409.
69. Lai, A., Clark, A.M., Escher, B.I., Fernandez, M., McEwen, L.R., Tian, Z., Wang, Z., and Schymanski, E.L. (2022). The Next Frontier of Environmental Unknowns: Substances of Unknown or Variable Composition, Complex Reaction Products, or Biological Materials (UVCBs). Environ. Sci. Technol. *56*, 7448–7466.
70. Wang, Z., Wiesinger, H., and Groh, K. (2021). Time to Reveal Chemical Identities of Polymers and UVCBs. Environ. Sci. Technol. *55*, 14473–14476.
71. Rajan, K., Zielesny, A., and Steinbeck, C. (2021). STOUT: SMILES to IUPAC names using neural machine translation. J. Cheminform. *13*. 34–14.
72. McKay, B.D., Yirik, M.A., and Steinbeck, C. (2022). Surge: a fast open-source chemical graph generator. J. Cheminform. *14*, 24–29.
73. Yirik, M.A., Sorokina, M., and Steinbeck, C. (2021). MAYGEN: an open-source chemical structure generator for constitutional isomers based on the orderly generation principle. J. Cheminform. *13*. 48–14.
74. Strynar, M., Dagnino, S., McMahen, R., Liang, S., Lindstrom, A., Andersen, E., McMillan, L., Thurman, M., Ferrer, I., and Ball, C. (2015). Identification of novel perfluoroalkyl ether carboxylic acids (PFECAs) and sulfonic acids (PFESAs) in natural waters using accurate mass time-of-flight mass spectrometry (TOFMS). Environ. Sci. Technol. *49*, 11622–11630.
75. Washington, J.W., Rosal, C.G., McCord, J.P., Strynar, M.J., Lindstrom, A.B., Bergman, E.L., Goodrow, S.M., Tadesse, H.K., Pilant, A.N., Washington, B.J., et al. (2020). Nontargeted mass-spectral detection of chloroperfluoropolyether carboxylates in New Jersey soils. Science *368*, 1103–1107.
76. Weber, E.J., Tebes-Stevens, C., Washington, J.W., and Gladstone, R. (2022). Development of a PFAS reaction library: identifying plausible transformation pathways in environmental and biological systems. Environ. Sci. Process. Impacts *24*, 689–753.
77. Washington, J.W., Jenkins, T.M., and Weber, E.J. (2015). Identification of unsaturated and 2H polyfluorocarboxylate homologous series and their detection in environmental samples and as polymer degradation products. Environ. Sci. Technol. *49*, 13256–13263.
78. Sleight, T.W., Khanna, V., Gilbertson, L.M., and Ng, C.A. (2020). Network analysis for prioritizing biodegradation metabolites of polycyclic aromatic hydrocarbons. Environ. Sci. Technol. *54*, 10735–10744.
79. NORMAN (2020). NORMAN Suspect List Exchange on Zenodo. NORMAN Suspect List Exchange: Zenodo Community. http://www.norman-network.com/?q=suspect-list-exchange.
80. Djoumbou-Feunang, Y., Fiamoncini, J., Gil-de-la-Fuente, A., Greiner, R., Manach, C., and Wishart, D.S. (2019). BioTransformer: a comprehensive computational tool for small molecule metabolism prediction and metabolite identification. J. Cheminform. *11*, 2–25.
81. Evich, M.G., Davis, M., Weber, E.J., Tebes-Stevens, C., Acrey, B., Henderson, W.M., Goodrow, S., Bergman, E., and Washington, J.W. (2022). Environmental Fate of Cl-PFPECAs: Predicting the Formation of PFAS Transformation Products in New Jersey Soils. Environ. Sci. Technol. *56*, 7779–7788.
82. Sigmund, G., Arp, H.P.H., Aumeier, B.M., Bucheli, T.D., Chefetz, B., Chen, W., Droge, S.T.J., Endo, S., Escher, B.I., Hale, S.E., et al. (2022). Sorption and Mobility of Charged Organic Compounds: How to Confront and Overcome Limitations in Their Assessment. Environ. Sci. Technol. *56*, 4702–4710.
83. Neuwald, I., Muschket, M., Zahn, D., Berger, U., Seiwert, B., Meier, T., Kuckelkorn, J., Strobel, C., Knepper, T.P., and Reemtsma, T. (2021). Filling the knowledge gap: A suspect screening study for 1310 potentially persistent and mobile chemicals with SFC-and HILIC-HRMS in two German river systems. Water Res. *204*, 117645.

84. Man, J., Zeng, L., Luo, J., Gao, W., and Yao, Y. (2021). Application of the Deep Learning Algorithm to Identify the Spatial Distribution of Heavy Metals at Contaminated Sites. ACS ES. T. Eng. *2*, 158–168.
85. Coley, C.W., Eyke, N.S., and Jensen, K.F. (2020). Autonomous discovery in the chemical sciences part I: Progress. Angew. Chem. Int. Ed. Engl. *59*, 22858–22893.
86. Coley, C.W., Eyke, N.S., and Jensen, K.F. (2020). Autonomous discovery in the chemical sciences part II: outlook. Angew. Chem. Int. Ed. Engl. *59*, 23414–23436.
87. Grimstein, M., Yang, Y., Zhang, X., Grillo, J., Huang, S.-M., Zineh, I., and Wang, Y. (2019). Physiologically based pharmacokinetic modeling in regulatory science: an update from the US Food and Drug Administration's Office of Clinical Pharmacology. J. Pharm. Sci. *108*, 21–25.
88. Steffen, W., Richardson, K., Rockström, J., Cornell, S.E., Fetzer, I., Bennett, E.M., Biggs, R., Carpenter, S.R., De Vries, W., De Wit, C.A., et al. (2015). Planetary boundaries: Guiding human development on a changing planet. Science *347*, 1259855.
89. Persson, L., Carney Almroth, B.M., Collins, C.D., Cornell, S., de Wit, C.A., Diamond, M.L., Fantke, P., Hassellöv, M., MacLeod, M., Ryberg, M.W., et al. (2022). Outside the safe operating space of the planetary boundary for novel entities. Environ. Sci. Technol. *56*, 1510–1521.
90. Lakshminarasimhan, R. (2022). Solar-driven water treatment: the path forward for the energy–water nexus. In Solar-Driven Water Treatment (Elsevier), pp. 337–362.
91. Cansino-Loeza, B., Munguía-López, A.d.C., and Ponce-Ortega, J.M. (2022). A water-energy-food security nexus framework based on optimal resource allocation. Environ. Sci. Pol. *133*, 1–16.
92. Alam, G., Ihsanullah, I., Naushad, M., and Sillanpää, M. (2022). Applications of artificial intelligence in water treatment for optimization and automation of adsorption processes: Recent advances and prospects. Chem. Eng. J. *427*, 130011.
93. Lohani, M.B., Singh, A., Rupainwar, D.C., and Dhar, D.N. (2008). Seasonal variations of heavy metal contamination in river Gomti of Lucknow city region. Environ. Monit. Assess. *147*, 253–263.
94. OECD (2012). OECD Studies on Water: A Framework for Financing Water Resources Management. https://www.oecd.org/env/resources/Financing%20WRM%20-In%20Brief%20-FINAL.pdf.
95. Scheffler, M., Aeschlimann, M., Albrecht, M., Bereau, T., Bungartz, H.-J., Felser, C., Greiner, M., Groß, A., Koch, C.T., Kremer, K., et al. (2022). FAIR data enabling new horizons for materials research. Nature *604*, 635–642.
96. Klemes, M.J., Skala, L.P., Ateia, M., Trang, B., Helbling, D.E., and Dichtel, W.R. (2020). Polymerized molecular receptors as adsorbents to remove micropollutants from water. Acc. Chem. Res. *53*, 2314–2324.
97. Bui, X.-T., Chiemchaisri, C., Fujioka, T., and Varjani, S. (2019). Introduction to recent advances in water and wastewater treatment technologies. Water and wastewater treatment technologies, 3–12.
98. Zhongming, Z., Linong, L., Xiaona, Y., Wangqiang, Z., and Wei, L. (2021). Toolkit for Water Policies and Governance. https://www.oecd.org/environment/toolkit-for-water-policies-and-governance-ed1a7936-en.htm.
99. Gingerich, D.B., and Mauter, M.S. (2018). Retrofitting the regulated power plant: optimizing energy allocation to electricity generation, water treatment, and carbon capture processes at coal-fired generating facilities. ACS Sustain. Chem. Eng. *6*, 2694–2703.
100. Clark, B., Gilles, G., and Tarpeh, W.A. (2022). Resin-Mediated pH Control of Metal-Loaded Ligand Exchangers for Selective Nitrogen Recovery from Wastewaters. ACS Appl. Mater. Interfaces *14*, 22950–22964.
101. Zhang, M., He, M., Chen, Q., Huang, Y., Zhang, C., Yue, C., Yang, L., and Mu, J. (2022). Feasible synthesis of a novel and low-cost seawater-modified biochar and its potential application in phosphate removal/recovery from wastewater. Sci. Total Environ. *824*, 153833.
102. Cho, S., Lim, J., Cho, H., Yoo, Y., Kang, D., and Kim, J. (2022). Novel process design of desalination wastewater recovery for CO2 and SOX utilization. Chem. Eng. J. *433*, 133602.
103. Taghvaie Nakhjiri, A., Sanaeepur, H., Ebadi Amooghin, A., and Shirazi, M.M.A. (2022). Recovery of precious metals from industrial wastewater towards resource recovery and environmental sustainability: A critical review. Desalination *527*, 115510.
104. Zhang, X., and Liu, Y. (2022). Circular economy is game-changing municipal wastewater treatment technology towards energy and carbon neutrality. Chem. Eng. J. *429*, 132114.
105. Rogers, M.D. (2003). Risk analysis under uncertainty, the Precautionary Principle, and the new EU chemicals strategy. Regul. Toxicol. Pharmacol. *37*, 370–381.
106. Schymanski, E.L., and Williams, A.J. (2017). Open science for identifying "known unknown" chemicals. Environ. Sci. Technol. *51*, 5357–5359.
107. Schymanski, E.L., and Bolton, E.E. (2021). FAIR chemical structures in the Journal of Cheminformatics. J. Cheminform. *13*, 50–53.
108. CODATA (2022). WorldFAIR: Global Cooperation on FAIR Data Policy and Practice – a Major Two-Year Project Starts Today. https://codata.org/worldfair-global-cooperation-on-fair-data-policy-and-practice-a-major-two-year-project-starts-today/.
109. US EPA (2022). New Chemicals Collaborative Research Program. https://www.epa.gov/reviewing-new-chemicals-under-toxic-substances-control-act-tsca/new-chemicals-collaborative.
110. European Commission (2011). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions Youth Opportunities Initiative. (European Commission Brussels, Belgium). URL: https://www.eea.europa.eu/policy-documents/communication-from-the-commission-to-1
111. Chemforward (2022). Apple, Google, and ChemFORWARD Join Forces to Drive Support for Safer Flame Retardants in the Electronics Sector - Website. https://www.chemforward.org/news/chemforward-apple-and-google-join-forces-to-drive-support-for-safer-flame-retardants-in-the-electronics-sector.
112. Blumenthal, J., Diamond, M.L., Hoffmann, M., and Wang, Z. (2022). Time to Break the "Lock-In" Impediments to Chemicals Management. Environ. Sci. Technol. *56*, 3863–3870.
113. European Commission (2020). Commission Implementing Regulation (EU) 2020/2087 of 14 December 2020 Concerning the Non-renewal of the Approval of the Active Substance Mancozeb. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32020R2087.
114. European Food Safety Authority (2018). Public Consultation on the Active Substance Tolpyralate (SL-573). https://www.efsa.europa.eu/en/consultations/call/180503.
115. U.S.Department of Agriculture (2020). European Union: Mancozeb Non-renewal and MRL Review. https://www.fas.usda.gov/data/european-union-mancozeb-non-renewal-and-mrl-review.
116. Hanisch, R., Chalk, S., Coulon, R., Cox, S., Emmerson, S., Flamenco Sandoval, F.J., Forbes, A., Frey, J., Hall, B., Hartshorn, R., et al. (2022). Stop squandering data: make units of measurement machine-readable. Nature *605*, 222–224.