Mohammed Taha *et al.*
Environmental Sciences Europe (2022) 34:104
https://doi.org/10.1186/s12302-022-00680-6

Environmental Sciences Europe

RESEARCH

Open Access

The NORMAN Suspect List Exchange (NORMAN-SLE): facilitating European and worldwide collaboration on suspect screening in high resolution mass spectrometry

Hiba Mohammed Taha[1], Reza Aalizadeh[2], Nikiforos Alygizakis[3,2], Jean-Philippe Antignac[4], Hans Peter H. Arp[5,6], Richard Bade[7], Nancy Baker[8], Lidia Belova[9], Lubertus Bijlsma[10], Evan E. Bolton[11], Werner Brack[12,13], Alberto Celma[10,14], Wen-Ling Chen[15], Tiejun Cheng[11], Parviel Chirsir[1], Ľuboš Čirka[16,3], Lisa A. D'Agostino[17], Yannick Djoumbou Feunang[18], Valeria Dulio[19], Stellan Fischer[20], Pablo Gago-Ferrero[21], Aikaterini Galani[2], Birgit Geueke[22], Natalia Głowacka[3], Juliane Glüge[23], Ksenia Groh[24], Sylvia Grosse[25], Peter Haglund[26], Pertti J. Hakkinen[11], Sarah E. Hale[5], Felix Hernandez[10], Elisabeth M.-L. Janssen[24], Tim Jonkers[27], Karin Kiefer[24], Michal Kirchner[28], Jan Koschorreck[29], Martin Krauss[12], Jessy Krier[1], Marja H. Lamoree[27], Marion Letzel[30], Thomas Letzel[31], Qingliang Li[11], James Little[32], Yanna Liu[33], David M. Lunderberg[34,35], Jonathan W. Martin[17], Andrew D. McEachran[36], John A. McLean[37], Christiane Meier[29], Jeroen Meijer[38], Frank Menger[14], Carla Merino[39,40], Jane Muncke[22], Matthias Muschket[12], Michael Neumann[29], Vanessa Neveu[41], Kelsey Ng[3,42], Herbert Oberacher[43], Jake O'Brien[7], Peter Oswald[3], Martina Oswaldova[3], Jaqueline A. Picache[37], Cristina Postigo[44,14], Noelia Ramirez[45,39], Thorsten Reemtsma[12], Justin Renaud[46], Pawel Rostkowski[47], Heinz Rüdel[48], Reza M. Salek[41], Saer Samanipour[49], Martin Scheringer[23,42], Ivo Schliebner[29], Wolfgang Schulz[50], Tobias Schulze[12], Manfred Sengl[30], Benjamin A. Shoemaker[11], Kerry Sims[51], Heinz Singer[24], Randolph R. Singh[1,52], Mark Sumarah[46], Paul A. Thiessen[11], Kevin V. Thomas[7], Sonia Torres[39], Xenia Trier[53], Annemarie P. van Wezel[54], Roel C. H. Vermeulen[38], Jelle J. Vlaanderen[38], Peter C. von der Ohe[29], Zhanyun Wang[55], Antony J. Williams[56], Egon L. Willighagen[57], David S. Wishart[58], Jian Zhang[11], Nikolaos S. Thomaidis[2], Juliane Hollender[23,24], Jaroslav Slobodnik[3] and Emma L. Schymanski[1*]

Abstract

Background: The NORMAN Association (https://www.norman-network.com/) initiated the NORMAN Suspect List Exchange (NORMAN-SLE; https://www.norman-network.com/nds/SLE/) in 2015, following the NORMAN collaborative trial on non-target screening of environmental water samples by mass spectrometry. Since then, this exchange

*Correspondence: emma.schymanski@uni.lu

[1] Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, 6 Avenue du Swing, 4367 Belvaux, Luxembourg
Full list of author information is available at the end of the article

of information on chemicals that are expected to occur in the environment, along with the accompanying expert knowledge and references, has become a valuable knowledge base for "suspect screening" lists. The NORMAN-SLE now serves as a FAIR (Findable, Accessible, Interoperable, Reusable) chemical information resource worldwide.

Results: The NORMAN-SLE contains 99 separate suspect list collections (as of May 2022) from over 70 contributors around the world, totalling over 100,000 unique substances. The substance classes include per- and polyfluoroalkyl substances (PFAS), pharmaceuticals, pesticides, natural toxins, high production volume substances covered under the European REACH regulation (EC: 1272/2008), priority contaminants of emerging concern (CECs) and regulatory lists from NORMAN partners. Several lists focus on transformation products (TPs) and complex features detected in the environment with various levels of provenance and structural information. Each list is available for separate download. The merged, curated collection is also available as the NORMAN Substance Database (NORMAN SusDat). Both the NORMAN-SLE and NORMAN SusDat are integrated within the NORMAN Database System (NDS). The individual NORMAN-SLE lists receive digital object identifiers (DOIs) and traceable versioning via a Zenodo community (https://zenodo.org/communities/norman-sle), with a total of > 40,000 unique views, > 50,000 unique downloads and 40 citations (May 2022). NORMAN-SLE content is progressively integrated into large open chemical databases such as PubChem (https://pubchem.ncbi.nlm.nih.gov/) and the US EPA's CompTox Chemicals Dashboard (https://comptox.epa.gov/dashboard/), enabling further access to these lists, along with the additional functionality and calculated properties these resources offer. PubChem has also integrated significant annotation content from the NORMAN-SLE, including a classification browser (https://pubchem.ncbi.nlm.nih.gov/classification/#hid=101).

Conclusions: The NORMAN-SLE offers a specialized service for hosting suspect screening lists of relevance for the environmental community in an open, FAIR manner that allows integration with other major chemical resources. These efforts foster the exchange of information between scientists and regulators, supporting the paradigm shift to the "one substance, one assessment" approach. New submissions are welcome via the contacts provided on the NORMAN-SLE website (https://www.norman-network.com/nds/SLE/).

Keywords: Suspect screening, High resolution mass spectrometry, Non-target screening, Open science, FAIR (Findable Accessible Interoperable Reusable) data, Data exchange, Cheminformatics, Exposomics, Environmental contaminants, Chemicals of emerging concern

Background

In environmental analytical chemistry, suspect screening typically involves the use of high resolution mass spectrometry (HRMS) to search for the presence of chemicals in environmental samples based on suspect lists, using the exact mass as a first step in the annotation of detected features [1, 2]. Suspect screening has grown in popularity over the last few years as an efficient way to complement traditional target analysis approaches, where a reference standard is required, without performing a time-intensive non-target screening of the tens of thousands of unknown features typical in environmental samples using extensive compound databases. Several publications describe these approaches in greater detail [1–4]. The NORMAN Association (a network of reference laboratories for monitoring of contaminants of emerging concern (CECs) in the environment—hereafter "NORMAN") [5] ran the first non-target screening (NTS) collaborative trial on river water in 2013/2014 [4]. The results showed that participants tentatively identified roughly as many chemicals via both suspect and target screening methods, but very few via NTS [4]. This early effort demonstrated that suspect screening approaches were more efficient and popular across the 19 participating institutes, offering a much higher annotation rate than non-target identification. Since then, NORMAN has run further collaborative trials involving suspect screening, including dust [6], passive samplers [7] and biota [8]. Suspect screening has also gained popularity beyond environmental studies and matrices, expanding recently to biomonitoring (e.g., [9, 10]).

One major outcome of the 2013/2014 NORMAN NTS collaborative trial was the clear need for a better exchange of chemical information both among and beyond NORMAN members [4], since the 2013/2014 collaborative trial participants used an incredibly wide variety of data sources during the trial (shown in Table 3 of [4]). This need had already been identified earlier, for example in the MODELKEY project [11] that included several NORMAN members, but the right implementation strategy remained elusive. A second NTS collaborative trial outcome, discussed in subsequent workshops, was a debate between "screen smart", versus "screen big". At the time, the "screen smart" strategy had been employed, for example, to study pesticides [12], pharmaceuticals [13] and surfactants [14] using relatively small lists (185, 980 and 394 entries, respectively), to support

focussed research questions. In contrast, the "screen big" strategy used very large lists containing thousands of chemicals (e.g., lists of high production volume chemicals registered under the European Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) regulation (EC No 1272/2008)) to find more hits—with the accompanying risk of many more false positives (see e.g., [15, 16]). Naturally, the boundary between these two strategies blurred over time, as some "smart" suspect lists also became quite "big". For instance, the STOFF-IDENT (https://water.for-ident.org/#!home) compilation of water-relevant contaminants such as pesticides, pharmaceuticals and industrial chemicals [17] includes over 10,500 substances. This list is "smart" with respect to the relevance to the water compartment, but with many pollutant classes and a large proportion of REACH chemicals, the overall number of chemicals is large enough to increase the probability of generating many false-positive results. In the extreme, "screen big" could be extended to candidates from even larger compound databases with millions of entries, which are commonly used in NTS approaches—with the lower success rates (i.e., more false positives) as mentioned above. Since suspect screening approaches typically start with only an exact mass of the expected adduct(s) of the suspects, there is a large burden of proof to confirm that the "suspect hit" is actually present, as discussed elsewhere [2–4].

The exchange of and access to chemical information in an open (i.e., free to access, publicly available) manner [18] has not always been as easy as it appears today. A key breakthrough was achieved in 2004 with the launch of PubChem (https://pubchem.ncbi.nlm.nih.gov/) [19], currently one of the largest open chemical knowledge bases with extensive information on over 111 million chemicals (July 2022). The ChemSpider collection was released a few years later (http://www.chemspider.com/) [20] and now contains 114 million chemicals (July 2022). The United States Environmental Protection Agency (US EPA) released the CompTox Chemicals Dashboard (https://comptox.epa.gov/dashboard/) [21] (hereafter "CompTox") in 2016 as a smaller collection, currently of 906,511 chemicals (July 2022) related to environmental and toxicology questions. Likewise, in 2016 the term "*FAIR*" was coined, describing how to make research more *F*indable, *A*ccessible, *I*nteroperable and *R*eusable [22, 23]. Together, ensuring that data is both Open and FAIR is a powerful combination [24]. The European Union (EU) is also embracing Open and FAIR principles. The European Chemicals Agency (ECHA) [25] and the European Food and Safety Authority (EFSA) [26] are transitioning their information to be more Open and FAIR, while Joint Research Centre (JRC) has released the Information Portal for Chemical Monitoring (IPCHEM) for the exchange of monitoring data in Europe [27]. Recent initiatives such as the European Partnership for Chemicals Risk Assessment (PARC) [28, 29] and the Environmental Exposure Assessment Research Infrastructure (EIRENE) [30] will strengthen this into the future.

In response to the NORMAN NTS collaborative trial outcomes, NORMAN initiated the NORMAN Suspect List Exchange (NORMAN-SLE, https://www.norman-network.com/nds/SLE/) in 2015 as part of the NORMAN Database System (NDS, https://www.norman-network.com/nds/) [29, 31] to facilitate the open access exchange of various suspect lists within and beyond Europe. This FAIR, open access, whole community initiative is not limited to NORMAN members. The primary aim of the NORMAN-SLE is to provide a location where suspect lists are publicly accessible, together with appropriate reference information, for interested parties to browse and select as desired (facilitating the "screen smart" approach). The NORMAN-SLE forms the basis for the NORMAN Substance Database (NORMAN SusDat, https://www.norman-network.com/nds/susdat/), a merged and curated data table with additional parameters for use in NORMAN activities (to facilitate the "screen big" approach), which will be described in more detail in a separate article. The present article covers the creation and implementation of the NORMAN-SLE as an Open and FAIR data resource, along with its integration with major open chemistry resources (PubChem, CompTox) as described below in the methods section, followed by an overview of the current state, implications and outlook in the results and discussion sections.

Methods

NORMAN Suspect List Exchange (NORMAN-SLE) website

The principle behind the NORMAN-SLE is simple: facilitating the exchange of chemical information to support the suspect screening of primarily organic contaminants amenable to liquid or gas chromatography (LC or GC) coupled to mass spectrometry. The website itself (https://www.norman-network.com/nds/SLE/) contains a simple overview of the background behind the NORMAN-SLE and a table containing the suspect lists themselves (with the fields "Number", "Abbreviation", "Description", "Link to full list", "Link to InChIKey list" and "References"), as shown in Fig. 1 and explained further below. Each list has a number (starting with S0 for SUSDAT, the merged collection), increasing sequentially with every contribution, along with an abbreviation for easier integration, access, and recognition.

The idea behind the simplicity of this website is to enable public access to various suspect lists as close as

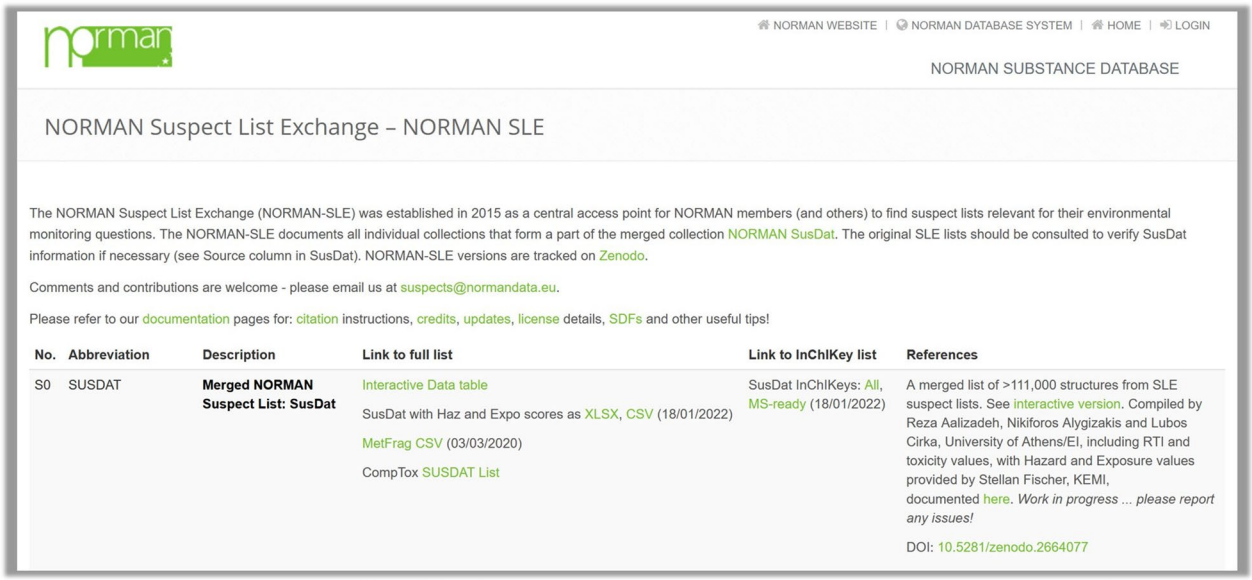

Fig. 1 Screenshot of the NORMAN Suspect List Exchange (https://www.norman-network.com/nds/SLE/) [32]

possible to the lists used in original publications, but with a reasonable degree of standardization and, where possible, added value to enhance and FAIRify these lists for future use (see below). If major adjustments were made to a submitted list, the original list is provided along with modified versions, so that both sets of information are available.

Information content and preparation of suspect lists hosted on the NORMAN-SLE

The minimum information available in most lists is a name and at least one additional identifier, although in most lists, far more information is available. At least one chemical name (plus other synonyms if available) should be included. The preferred formats for structural information are the simplified molecular-input line-entry system (SMILES) [33] plus the International Chemical Identifier (InChI) in the form of standard InChI and InChIKey [34]. Common database identifiers provided typically include one (or more) of either Chemical Abstract Service (CAS) number(s) [35], EC number [36], PubChem Compound Identifier (CID) [19], ChemSpider identifier (CSID) [20] and/or the Distributed Structure-Searchable Toxicity (DSSTox) substance identifier (DTXSID) used in CompTox [21]. To support suspect screening, the (neutral) monoisotopic masses and molecular formulae are included in many of the lists. This information, along with several other predicted values, is also included in the merged NORMAN SusDat. Several other fields may be present, depending on the context of the suspect list, and are included where available. More details on the chemical structure identifiers and recommended chemical structural data templates are provided elsewhere [24, 37].

The suspect lists (commonly submitted via email to NORMAN contact points, see Fig. 2, top left) are processed upon submission, with the subsequent processing steps highly dependent on both the type of submission and the size of the list. While the suspect list number is assigned sequentially, the abbreviation, name and description are assigned following pre-defined conventions, and in discussion with authors. Where necessary, curation is performed on these lists to fill in missing values where at least a chemical identifier and/or structural information and/or (correct) name was provided. For some lists, the missing values are filled using automated workflows covering a variety of web services (depending on the list and contributor) from PubChem [19], ChemSpider [20] and CACTUS (https://cactus.nci.nih.gov/), typically via RMassBank [38], RChemMass [39] and other related packages in the R programming language. Other lists are processed with batch services offered through PubChem [19, 40] and CompTox [21, 41]. Additional chemical structure interconversions (e.g., SMILES to InChI) are performed with OpenBabel (http://openbabel.org/) [42] or the Chemistry Development Kit (CDK) (usually via R) [43] where necessary. Note that the curation performed on the individual suspect lists is independent of the curation and merging to form the NORMAN SusDat collection (see Fig. 2, bottom left), which will be detailed in a separate publication. The processes evolve over time as new technical possibilities

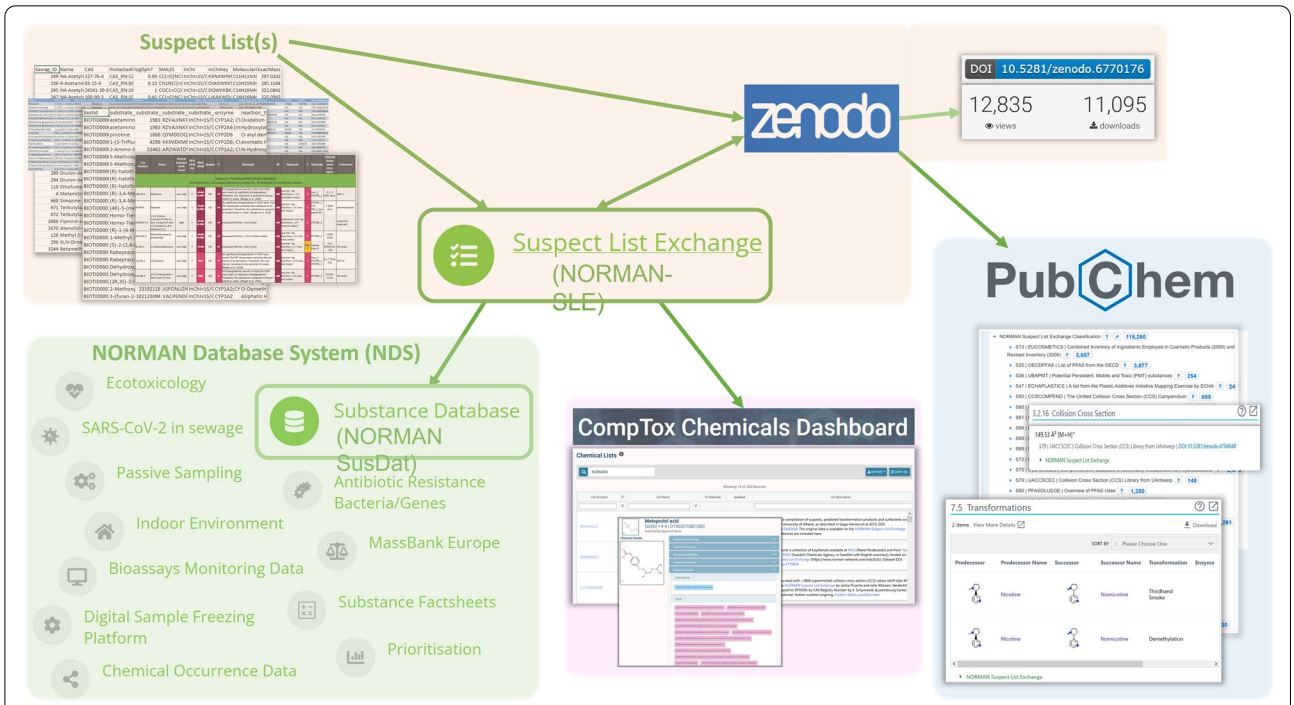

Fig. 2 Schematic showing the relationships between submitted suspect lists, the NORMAN-SLE and downstream resources. Top (orange shading): suspect lists submitted in various formats are curated, then added to the NORMAN-SLE website (centre) and archived on the NORMAN-SLE Zenodo community (top right), yielding a DOI and use statistics. Bottom left (green shading): the NORMAN-SLE serves as an information source for NORMAN SusDat and the NORMAN Database System (NDS). Bottom middle (pink shading): NORMAN-SLE lists are integrated in CompTox manually. Bottom right (blue shading): NORMAN-SLE content is harvested from Zenodo via mapping files and integrated into PubChem in an automated workflow

arise (e.g., batch searching). The resulting suspect lists are generally provided as Excel (XLSX) and comma separated values (CSV) formats, as standardized as reasonably possible, on the website. The CSV format provides greater interoperability, including allowing import into various libraries, vendor and open software, as well as PubChem (described below). A separate InChIKey file is also provided, as this allows fast screening of suspects within the in silico fragmenter MetFrag [44] and other approaches. For some of the lists, additional files are provided, to disseminate all the relevant details. Finally, references and additional information are given, to acknowledge contributors, but also to provide users quick access to the rationale behind each individual suspect list. Further details on the NORMAN-SLE contents, including references, are given in the Results section.

Several suspect lists contain partial, incomplete, or even no structural information, such as the per- and polyfluoroalkyl substances (PFAS) lists S9 PFASTRIER [45] (e.g., elemental compositions retrieved from patents where no structural or isomer information was available) and S46 PFASNTREV19 [46, 47] (a compilation of PFAS identification efforts in non-target screening studies), as well as the surfactant isomer list S18 TSCASURF [48]. Nevertheless, these lists still provide vital information for identification by mass and/or molecular formula (see e.g., [14, 49], where whole surfactant classes can be identified via the general formula of a homologous series of several structural isomers). For those lists with partial information, missing values were filled in, where possible, as described above, and were saved in separate files or as multiple sheets in one file. Associated InChIKey lists were only generated for known structures. Dealing with partially characterized molecular features or chemical substances of Unknown or Variable Composition, Complex Reaction Products or Biological Materials (UVCB substances, UVCBs) is a subject of future collaborations beyond the scope of the current article (see e.g., [50, 51]), as discussed further below.

NORMAN-SLE on Zenodo

The development of the Zenodo repository [52] enabled public archiving, versioning and generation of a Digital Object Identifier (DOI) for each NORMAN-SLE list. Thus, since 2019, the NORMAN-SLE content has been uploaded to and archived on the Zenodo repository [52],

gathered under the NORMAN-SLE community (https://zenodo.org/communities/norman-sle/) [53]. Each individual NORMAN-SLE collection has its own Zenodo record and thus a dataset DOI, allowing users to cite the individual lists directly, including specific versions, or all versions. Updates to lists can thus be tracked under the Zenodo version control system, with the master DOI always redirecting to the latest version. The lists are tracked under a versioning system following the pattern NORMAN-SLE-SXX-0.Y.Z, where SXX refers to the list number (as on the NORMAN-SLE website and as described below) and the 0.Y.Z pattern records whether it was a major update (Y is increased incrementally by 1) or minor update (Z is increased incrementally by 1). The leading "0" is currently a buffer. Major updates constitute new entries (e.g., new chemicals, rows, information, updates) to the lists, while minor updates are corrections or adjustments to the current contents without adding major new content (e.g., correcting names, identifiers, typographical errors). The presence on Zenodo has enabled better citation, the tracking of use statistics at an individual list level and additional possibilities for the integration with external resources such as PubChem, as shown in Fig. 2 (right) and discussed further below. Figure 3 shows the presence of the NORMAN-SLE on Zenodo, including versioning in the inset.

NORMAN-SLE and CompTox Chemicals Dashboard integration

Since CompTox [21] is a highly relevant resource for environmental and toxicological information, integration of NORMAN-SLE content is of interest to both parties and is achieved via the "Chemical Lists" functionality (https://comptox.epa.gov/dashboard/chemical-lists/). The integration started in 2017 and is performed through the upload of the DTXSIDs associated with the individual NORMAN-SLE lists to the DSSTox database [55] that underlies CompTox. Most lists have the NORMAN keyword associated with it, such that they are accessible through the URL https://comptox.epa.gov/dashboard/chemical-lists?search=NORMAN, or through a direct URL composed of the list code (e.g., https://comptox.epa.gov/dashboard/chemical-lists/BISPHENOLS for the S20 BISPHENOLS list). Several lists on the NORMAN-SLE were produced in a collaborative curation effort (e.g., S24 HUMANNEUROTOX [56], S37 LITMINEDNEURO [57] and S43 NEUROTOXINS [58], as part of [59]), or were curated and registered by the DSSTox curation team before uploading to the SLE (e.g., S25 OECDPFAS [60–62]). Some other lists on the NORMAN-SLE were sourced directly from CompTox as they contained entries highly relevant for the NORMAN Database System (e.g., S45 SYNTHCANNAB [63] and S58 PSYCHOCANNAB

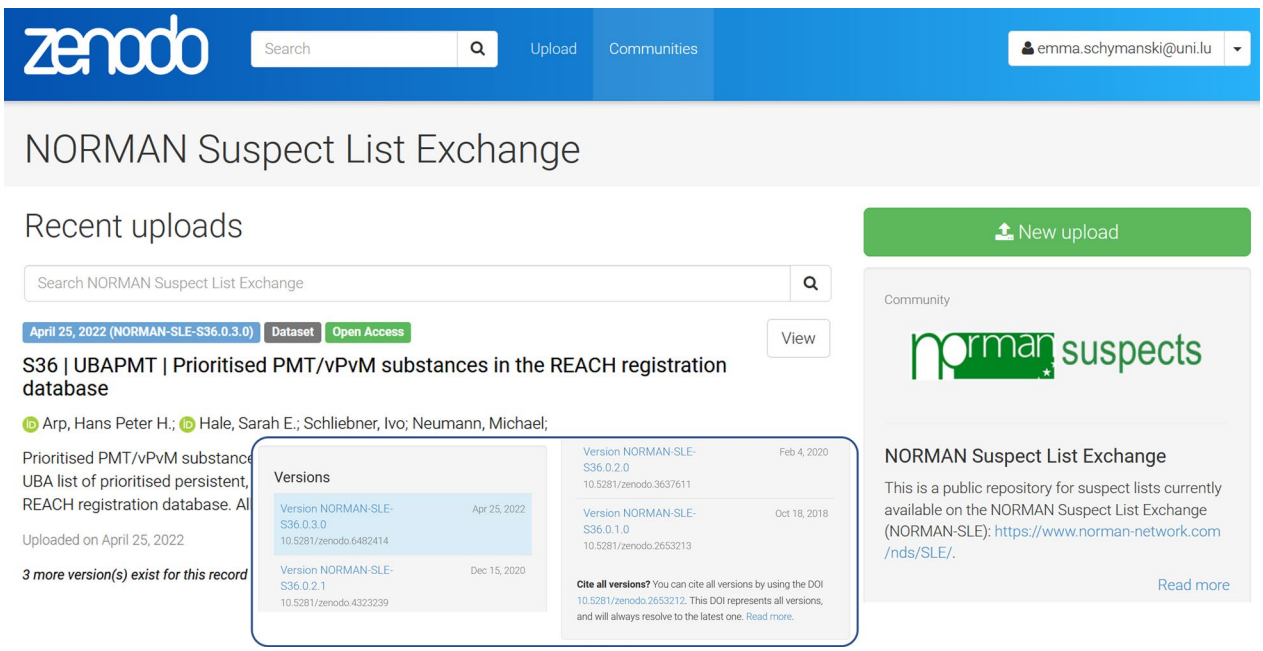

Fig. 3 The NORMAN Suspect List Exchange Zenodo community (https://zenodo.org/communities/norman-sle) with inset showing the versioning history of S36 UBAPMT (https://doi.org/10.5281/zenodo.2653212) [53, 54]

[64]). For recent lists, generally the CompTox batch search (https://comptox.epa.gov/dashboard/batch-search) [65] is used to retrieve DTXSIDs on the basis of the user-provided information, which are then provided directly to CompTox along with the list code, name and description for upload. The presence of compounds in NORMAN-SLE lists appear on the individual chemical records in CompTox (see pink entries in the inset in Fig. 2) and can also be identified by prefiltering in the CompTox batch search interface and including flags in the export files.

Due to the infrequent release of updates to CompTox, it may be many weeks or months before new NORMAN-SLE lists are available publicly on CompTox. Currently, 88 of the 99 NORMAN-SLE lists are on CompTox (see Additional file 1), with 74 listed under the "NORMAN" URL above. Since not all substances in the NORMAN-SLE are currently present in CompTox, the mapping of NORMAN-SLE lists in CompTox is often incomplete, i.e., the lists on CompTox contain only entries for which DTXSIDs currently exist (further details are provided in Additional file 1).

NORMAN-SLE and PubChem integration

As one of the largest open chemical databases with millions of monthly users, integration of NORMAN-SLE content in PubChem has great potential to increase the visibility of this community effort. The NORMAN-SLE integration with PubChem [19] (https://pubchem.ncbi.nlm.nih.gov/) commenced in 2019. The first substance deposition was processed on November 22, 2019. The deposition file is compiled from all lists by the PubChem team, via a mapping file hosted on the Environmental Cheminformatics (ECI) group (University of Luxembourg) GitLab pages [66]. This mapping file contains a link to the latest version of each suspect list (CSV file) on Zenodo, the list details and version, the dataset DOI, extra DOIs (to include related publications), mappings to the columns containing the chemical identifiers (SMILES, InChIKey, InChI, Synonym), the NORMAN-SLE URL and a comment field. The compiled deposition file is mapped to PubChem Substance Identifiers (SIDs) and PubChem Compound Identifiers (CIDs) via the PubChem deposition system. While SIDs are available for all substances deposited to PubChem (including those with undefined structures), CIDs are only available for all unique chemical structures (i.e., defined chemical structures) extracted from substance depositions via the PubChem standardization process [67]. As a result, the number of compounds (CIDs) will generally be less than the number of substances (SIDs). Any SMILES errors found during deposition are debugged in collaboration with the PubChem team and any dataset-specific causes are fixed in the corresponding NORMAN-SLE datasets by releasing new minor versions on Zenodo (see e.g., descriptions in [68, 69]). Synonyms are currently provided as a small, manually curated file containing the columns CID, InChIKey, Synonym, Reference DOI and Dataset information (114 entries on 30 April 2022, see [70]) to specifically add missing synonyms to PubChem [70]. These are primarily newly deposited structures (i.e., structures not yet in PubChem) associated with S74 REFTPS [71] and S96 ECIPFAS [72]. The PubChem/NORMAN-SLE deposition is re-run once updates are available and takes minutes to run. The updated data are live on the public PubChem website within hours to days (newly added structures can take longer to index fully). The latest deposition and number of live substances (i.e., the number of substances currently available on the public website) can be retrieved from the NORMAN-SLE data source page in PubChem [73].

The contents of individual NORMAN-SLE lists are available interactively in PubChem via the NORMAN Suspect List Exchange Tree (https://pubchem.ncbi.nlm.nih.gov/classification/#hid=101, hereafter "PubChem NORMAN-SLE Tree") on the PubChem Classification Browser [74]. This is compiled by PubChem from a second mapping file, also hosted on the ECI GitLab pages [75]. For each dataset, this mapping file contains a link to the latest InChIKey file on Zenodo, the list title as it should appear in the tree (e.g., "`S00 | SUSDAT | Merged NORMAN Suspect List: SusDat`") and a tool tip, i.e., further details about the list that displays when users click the "?" icon on the Classification Browser (see figure in Results section). The mapping file also contains additional fields defining the content of interest (keywords, annotations) and other information for internal housekeeping. All lists (except S18 TSCASURF, for which no InChIKeys exist) are listed in numerical order in the PubChem NORMAN-SLE Tree. In addition, certain lists with detailed classification content appear again at the top of the browser. These are mapped via structural information in the CSVs (not the InChIKey files) to profit from the detailed additional information available in these lists. The PubChem Classification Browser can also be accessed programmatically (i.e., in an automated manner), with documentation available on PubChem [67] and the ECI GitLab pages [76]. The PubChem NORMAN-SLE Tree also enables users to download individual lists (or even various combinations thereof via advanced queries) in the variety of formats offered by PubChem, including the structure data format (SDF) not currently offered on the NORMAN-SLE website, see documentation available in e.g., [77].

PubChem has also integrated several categories of annotation content, i.e., detailed information about individual chemicals, into the compound records in PubChem. As of

30 April 2022, a total of 17 annotation categories, which equate to headers in the Table of Contents entries in PubChem [78], were integrated. Many relate to the chemical role or use (e.g., the Anatomical Therapeutic Chemical (ATC) Code for pharmaceuticals, Agrochemical Category, Chemical Classes, Use Classifications and Uses) and transformation information (e.g., included in the Transformations, Metabolism/Metabolites, Drug Transformations and Agrochemical Transformations headers). Others relate to chemical information (e.g., molecular formula) and measurement data, such as nuclear magnetic resonance (NMR—^{13}C, ^{19}F, ^{1}H, and ^{31}P), tandem MS (MS/MS) data and collision cross section (CCS) data from ion mobility experiments. Finally, taxonomy information (functionality recently added to PubChem [79] for organisms) has been included for some lists. All files necessary for the integration of the annotation content within PubChem are present in the Zenodo repository for the respective list, supported by additional mapping or annotation files either added in Zenodo, or hosted on the ECI GitLab pages in the "annotations" subfolder [80] where necessary. The latest overview and the entire content integrated in PubChem (in JSON, XML and ASNT formats, accessible programmatically or for download) is available from the NORMAN-SLE data source page in PubChem [73].

Results

Overview of NORMAN-SLE

The NORMAN-SLE includes 99 contributions (starting at S0 SUSDAT, the compilation of all NORMAN-SLE lists, to S98 TIRECHEM) from over 70 contributors as of May 2022, summarized in Fig. 4 and Table 1. Full details on all lists are available in Additional file 1 [81], including list details and chemical numbers across the resources in CSV format, and Additional file 2 [82], a May 2022 copy of the NORMAN-SLE website contents.

Figure 4 and Table 1 show the number of entries in each NORMAN-SLE list as present on the NORMAN-SLE website and in the latest versions on the NORMAN-SLE Zenodo collection as of May 2022. The number of InChIKeys associated with these lists (as of May 2022) is available in Additional file 1 [81]. Additional file 1 also includes the number of entries included in PubChem (obtained via the PubChem NORMAN-SLE Tree [74]) and CompTox (via both the CompTox Chemical Lists [232] website as well as via the PubChem EPA DSSTox Tree [233], since the latter can be automated). These statistics were compiled on 4 May 2022. The corresponding files and code are available at the ECI NORMAN-SLE GitLab repository [234] in the "stats" subfolder. Note that the addition of new content to the NORMAN-SLE was put on hold during compilation of this manuscript (May and June 2022), to ensure that the results included here are internally consistent. All statistics presented here reflect the data in this state. Updates resumed 28 June 2022 and will be described in later efforts (see "Future updates" below).

Summary statistics of the NORMAN-SLE

A selection of summary statistics and facts for the NORMAN-SLE is given in Table 2. Both the list and citation information were summarized on 4 May 2022 and the NORMAN-SLE PubChem numbers on 12 May 2022. The (cumulative) numbers of unique views and downloads collected from the NORMAN-SLE Zenodo community on 28 April 2022 are summarized in Table 3, along with the citation numbers for all lists and for the 5 most popular lists according to unique views. The "total unique compounds" number indicates how many entries have a defined chemical structure in PubChem, i.e., a PubChem CID. The "total live substances" number indicates how many entries are deposited, i.e., with a PubChem SID. The total number of unique compounds in PubChem is currently larger than S0 SUSDAT due to the different timing associated with the release cycle of NORMAN SusDat (the basis for S0 SUSDAT), as well as differences in the mappings of structures to unique identifiers. Future efforts will aim to close this time gap between NORMAN-SLE and NORMAN SusDat (see "Future updates" below). The data files supporting these statistics, including a breakdown of the DOIs of the citing articles, are archived on the ECI NORMAN-SLE GitLab pages [234] ("stats" subfolder) and are available as Additional file 3 [235] (views, downloads, citations per list) and Additional file 4 [236] (more detailed citation breakdown).

In total, 24 of the SLE lists have citations listed in Zenodo, with 40 citations from 19 articles. A full breakdown is given in Additional file 4 [236]. Of these 19 articles, 12 can be considered "internal", i.e., articles written by authors involved with the NORMAN-SLE, including 5 articles describing SLE datasets [59, 118, 149, 154, 174] and 7 others citing SLE lists [24, 142, 237–241], while 7 articles are external [242–248]. Of the 24 lists cited, 6 lists are cited by external authors: S0 SUSDAT, S13 EUCOSMETICS, S14 KEMIPFAS, S25 OECDPFAS, S46 NTPFASREV19 and S75 CyanoMetDB.

NORMAN-SLE PubChem integration

As described above, the NORMAN-SLE content has been integrated into PubChem in a variety of ways. The basis of all further integration is the substance depositions, formed from the compilation of all lists as described in the Methods section. As of 12 May 2022, the substance deposition in PubChem included 117,071 substances

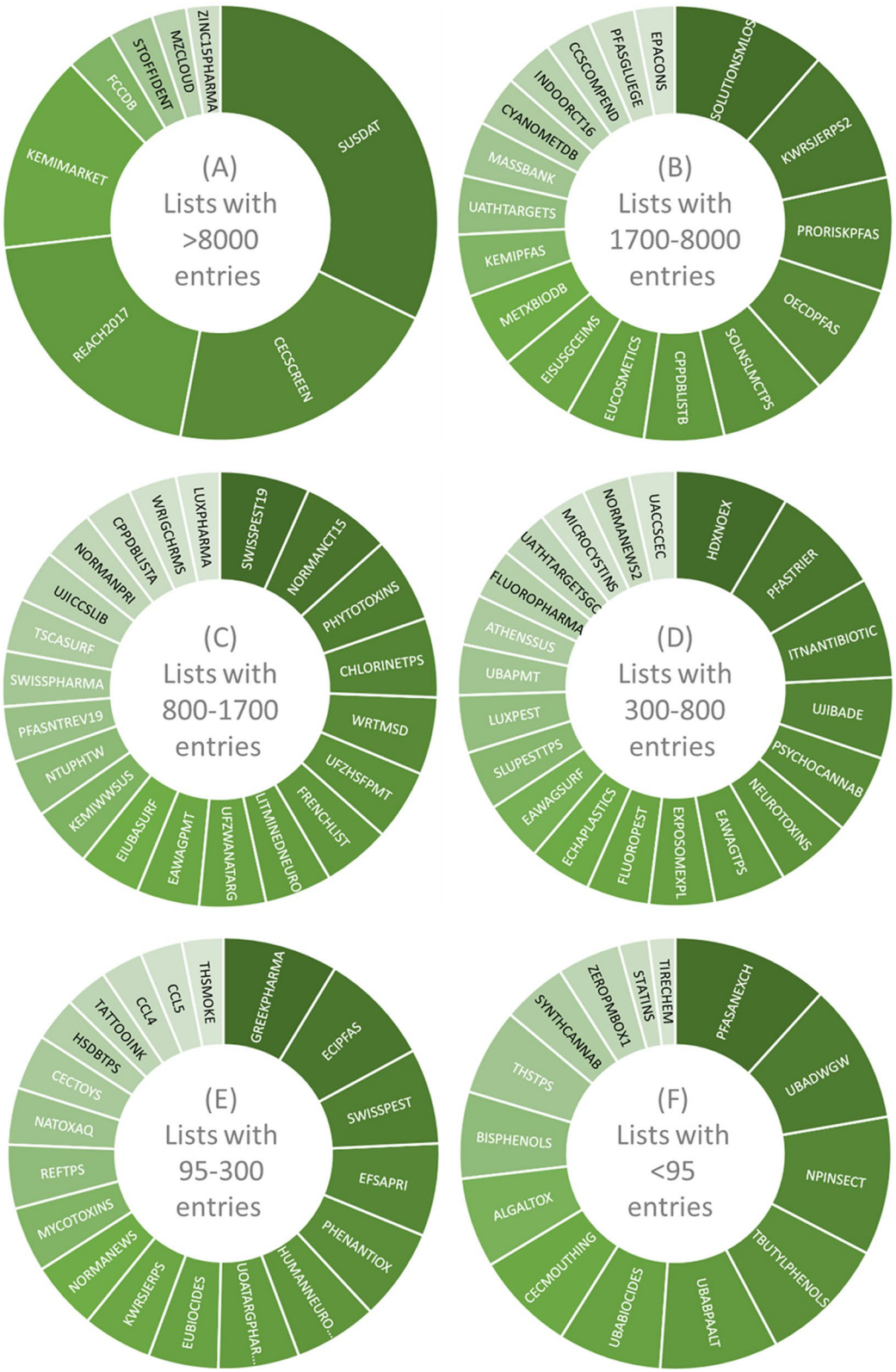

Fig. 4 Starburst plots of the 99 suspect lists forming the NORMAN-SLE contents. Lists with: (**A**) > 8000 entries; (**B**) 1700–8000 entries; (**C**) 800–1700 entries; (**D**) 300–800 entries; (**E**) 95–300 entries and (**F**) < 95 entries (ranges chosen to optimize plotting). The list codes, numbers of chemical entries and references are summarized in Table 1 according to the same groups, with full details in Additional file 1

Table 1 Summary of the NORMAN-SLE datasets, split by the groups shown in Fig. 4, with suspect list number (S), code, number of entries (lines in the file, in italics) and the accompanying references

Group	List number, Code, Entries (total lines) and References
(A) > 8000	S0 SUSDAT *109,631* [69, 83]; S71 CECSCREEN *70,397* [84, 85]; S32 REACH2017 *68,679* [86]; S17 KEMIMARKET *50,308* [68]; S77 FCCDB *12,285* [87–89]; S2 STOFFIDENT *11,289* [17, 90]; S19 MZCLOUD *8742* [91]; S55 ZINC15PHARMA *8646* [92–94]
(B) 1700–8000	S33 SOLUTIONSMLOS *6463* [95–97]; S27 KWRSJERPS2 *5702* [98, 99]; S89 PRORISKPFAS *4777* [100, 101]; S25 OECDPFAS *4725* [60–62]; S38 SOLNSLMCTPS *4465* [96, 97, 102]; S49 CPPDBLISTB *3353* [103–105]; S13 EUCOSMETICS *3333* [106–108]; S70 EISUSGCEIMS *3266* [109]; S73 METXBIODB *3148* [110, 111]; S14 KEMIPFAS *2602* [112, 113]; S21 UATHTARGETS *2466* [114, 115]; S1 MASSBANK *2305* [38, 116, 117]; S75 CYANOMETDB *2124* [118, 119]; S35 INDOORCT16 *2056* [6, 120]; S50 CCSCOMPEND *1983* [121–123]; S80 PFASGLUEGE *1926* [124, 125]; S22 EPACONS *1705* [126, 127]
(C) 800–1700	S60 SWISSPEST19 *1664* [128, 129]; S3 NORMANCT15 *1662* [4, 130]; S29 PHYTOTOXINS *1586* [131, 132]; S87 CHLORINETPS *1470* [133, 134]; S31 WRTMSD *1429* [135, 136]; S84 UFZHSFPMT *1310* [137–139]; S16 FRENCHLIST *1256* [140]; S37 LITMINEDNEURO *1243* [57, 59]; S53 UFZWANATARG *1235* [141]; S82 EAWAGPMT *1162* [142, 143]; S23 EIUBASURF *1154* [144]; S39 KEMIWWSUS *1123* [145]; S72 NTUPHTW *1068* [146, 147]; S46 PFASNTREV19 *1030* [46, 47]; S10 SWISSPHARMA *1020* [13, 148]; S18 TSCASURF *985* [48]; S61 UJICCSLIB *970* [149, 150]; S15 NORMANPRI *967* [151]; S48 CPPDBLISTA *902* [103, 105, 152]; S51 WRIGCHRMS *892* [153]; S76 LUXPHARMA *816* [154, 155]
(D) 300–800	S42 HDXNOEX *765* [156, 157]; S9 PFASTRIER *746* [45]; S6 ITNANTIBIOTIC *676* [158, 159]; S4 UJIBADE *544* [160, 161]; S58 PSYCHOCANNAB *531* [64]; S43 NEUROTOXINS *511* [58, 59]; S66 EAWAGTPS *486* [162, 163]; S34 EXPOSOMEXPL *440* [164–166]; S94 FLUOROPEST *423* [167, 168]; S47 ECHAPLASTICS *418* [169, 170]; S7 EAWAGSURF *410* [14, 171]; S78 SLUPESTTPS *400* [172, 173]; S69 LUXPEST *386* [174, 175]; S36 UBAPMT *341* [54, 176, 177]; S8 ATHENSSUS *340* [49, 178]; S92 FLUOROPHARMA *340* [179, 180]; S65 UATHTARGETSGC *334* [181, 182]; S62 NORMANEWS2 *321* [183, 184]; S85 MICROCYSTINS *321* [118, 185]; S79 UACCSCEC *311* [186, 187]
(E) 95–300	S57 GREEKPHARMA *263* [188]; S96 ECIPFAS *258* [72]; S11 SWISSPEST *218* [12, 189]; S54 EFSAPRI *212* [190, 191]; S30 PHENANTIOX *209* [192]; S24 HUMANNEUROTOX *190* [56, 59]; S56 UOATARGPHARMA *185* [193–195]; S28 EUBIOCIDES *160* [196]; S5 KWRSJERPS *159* [99, 197]; S12 NORMANEWS *156* [198, 199]; S26 MYCOTOXINS *149* [200]; S74 REFTPS *146* [71]; S64 NATOXAQ *130* [201, 202]; S91 CECTOYS *126* [203, 204]; S68 HSDBTPS *101* [174, 205]; S86 TATTOOINK *98* [206–208]; S41 CCL4 *96* [209, 210]; S83 CCL5 *96* [211, 212]; S52 THSMOKE *95* [213]
(F) < 95	S95 PFASANEXCH *94* [214, 215]; S63 UBADWGW *84* [176, 216]; S59 NPINSECT *83* [217]; S67 TBUTYLPHENOLS *77* [218]; S97 UBABPAALT *71* [219, 220]; S88 UBABIOCIDES *62* [221–223]; S93 CECMOUTHING *60* [203, 204]; S40 ALGALTOX *54* [224]; S20 BISPHENOLS *52* [225, 226]; S81 THSTPS *52* [227]; S45 SYNTHCANNAB *39* [63]; S90 ZEROPMBOX1 *38* [228, 229]; S44 STATINS *18* [230]; S98 TIRECHEM *16* [231]

Full details given in Additional file 1 and Additional file 2 [81, 82]

(i.e., with PubChem SIDs), mapping to 115,248 unique PubChem CIDs according to the compiled CIDs at the top of the PubChem NORMAN-SLE Tree [74] (see also Table 2). All lists except S18 TSCASURF (for which no InChIKeys are available) are included in the numerically ordered set of lists on the PubChem NORMAN-SLE Tree. As of 30 April 2022, additional detailed classification breakdowns were available for S13 EUCOSMETICS [108], S25 OECDPFAS [60], S36 UBAPMT [54], S47 ECHAPLASTICS [170], S50 CCSCOMPEND [121], S60 SWISSPEST19 [129], S61 UJICCSLIB [150], S66 EAWAGTPS [163], S68 HSDBTPS [205], S69 LUXPEST [175], S72 NTUPHTW [147], S75 CYANOMETDB [119], S79 UACCSCEC [187] and S80 PFASGLUEGE [124]. Detailed classification content for S77 FCCDB [89] is already drafted on the test site. A screenshot of the top portion of the PubChem NORMAN-SLE Tree is shown on the left in Fig. 5. The collision cross section (CCS) content (S50 CCSCOMPEND [121], S61 UJICCSLIB [150] and S79 UACCSCEC [187]) has also been merged and extended in the "Aggregated CCS Classification" tree on PubChem to combine this with the data from CCSbase [249, 250] and to allow browsing by adduct categories across all datasets [251]. All datasets mentioned here can be accessed via hyperlinks available at the NORMAN-SLE Data Source page on PubChem [73]. Documentation

Table 2 Selected overall summary statistics for the NORMAN-SLE, compiled in May 2022

Category	Number	Comment
Total number of lists	99	S0 to S98
Total unique compounds	115,248	From PubChem NORMAN-SLE Tree [74]
Total live substances	117,071	From PubChem NORMAN-SLE Data Source Page [73]
Total live annotations	21,114	From PubChem NORMAN-SLE Data Source Page [73]
Largest list (# entries)	109,631	S0 SUSDAT
Smallest list (# entries)	16	S98 TIRECHEM
Total list citations	40	From NORMAN-SLE Zenodo Community [53]

Further details are given in the "stats" subfolder of the ECI NORMAN-SLE GitLab repository [234]. Total unique compounds = CID count; total live substances = SID count, # entries = number of entries (i.e., rows) in the SLE lists

Table 3 Unique views, downloads and citations for all NORMAN-SLE lists and the Top 5 lists (by unique views), according to the NORMAN-SLE Zenodo Community [53]

List	Code	Unique views	Unique downloads	Citations
Top 5 Lists (sorted by unique views)				
S13	EUCOSMETICS [108]: Cosmetics	10,429	9088	2
S60	SWISSPEST19 [129]: Pesticides	2440	2316	3
S72	NTUPHTW [147]: Pharmaceuticals	2278	2083	0
S73	METXBIODB [110]: BioTransformer data	2043	503	2
S0	SUSDAT [69]: Merged database	1625	1858	6
Total values				
All	Totals over all lists	42,358	53,651	40

Statistics compiled on 28 April (views/downloads) and 5 May (citations) 2022. The corresponding raw data are given in Additional file 3: Table S3 [235] and on the ECI NORMAN-SLE GitLab pages [234]

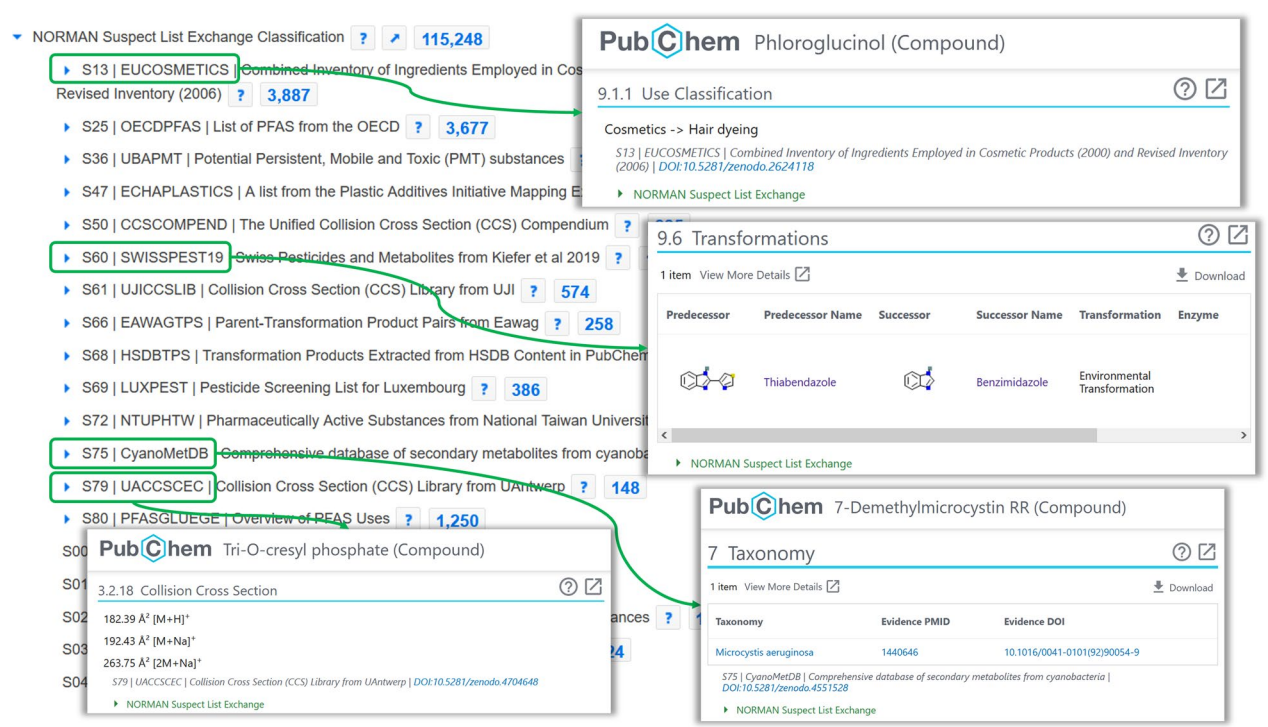

Fig. 5 A collage of NORMAN-SLE content in PubChem. Left/back: the PubChem NORMAN-SLE Tree, with entries containing detailed classifications at the top, indicated by the blue arrows. Insets: selected annotation content (Use Classification, Transformations, Taxonomy and Collision Cross Section), linked to the corresponding source list via the green boxes and arrows. Screenshots taken 30 May 2022 (taxonomy on 16 June 2022)

on how to access the data integrated within PubChem is provided on the ECI GitLab pages, including how to find MS [252] and CCS [253] data for NORMAN-SLE lists via PubChem. This also includes code to retrieve the CCS data [254], along with a compiled archive of all CCS values in PubChem (7 June 2022) on Zenodo [255].

In addition to the deposition and classification, extensive annotation content (i.e., expert knowledge) provided within the NORMAN-SLE lists has been integrated within PubChem. Various pieces of information from NORMAN-SLE lists now appear on the individual compound records for 21,114 compounds (12 May 2022), with several examples shown as insets in Fig. 5. While the presence of this annotation information in text form in individual PubChem records is useful for readers of the individual chemical records, it also helps in search engine optimization (SEO), i.e., the discovery of this information in generalized search engines, beyond the original database. Some categories (PubChem headings indicated in italics) relate to the chemical role, e.g., the "*ATC Code*"

for pharmaceuticals (from S66 EAWAGTPS [163] and S76 LUXPHARMA [155]), "*Agrochemical Category*" (S66 EAWAGTPS [163] and S69 LUXPEST [175]), or "*Chemical Classes*" (S75 CyanoMetDB [119]). Information in the "*Use Classifications*" and "*Uses*" categories come from S13 EUCOSMETICS [108], S25 OECDPFAS [60], S47 ECHAPLASTICS [170], S60 SWISSPEST19 [129], S66 EAWAGTPS [163], S69 LUXPEST [175], S72 NTUPHTW [147], S79 UACCSCEC [187] and S80 PFASGLUEGE [124]. The composite "*Molecular Formula*" representation in S80 PFASGLUEGE [124] is also integrated. Taxonomy information (for organisms) has been included under the "*Taxonomy*" heading for compounds present in S75 CyanoMetDB [118, 119] and S29 PHYTOTOXINS [132] from the Toxic Plants-Phytotoxins database [131], and also appears on the individual taxa pages.

Transformations for 5135 CIDs have been added from the datasets S60 SWISSPEST19 [129], S66 EAWAGTPS [163], S68 HSDBTPS [205], S73 METXBIODB [110], S74 REFTPS [71], S78 SLUPESTTPS [173] and S79 UACCSCEC [187], as described in some of the articles mentioned above [24, 174, 241]. As a part of this, SEO text snippets describing these relationships have been added to the following headings: *Metabolism/Metabolites* (S73 METXBIODB [110] and S82 THSTPS [227]), *Drug Transformations* (S66 EAWAGTPS [163]) and *Agrochemical Transformations* (S60 SWISSPEST19 [129], S66 EAWAGTPS [163] and S78 SLUPESTTPS [173]). An example *Transformations* entry is provided in the middle right inset in Fig. 5. The *Transformations* data are compiled and archived on GitLab [80] and Zenodo [256], and is integrated in patRoon 2.0 [257], an open source software for mass spectrometry based non-target analysis that includes suspect and transformation product screening workflows.

Finally, a significant amount of experimental data has also been included in PubChem from NORMAN-SLE contributors. MS/MS and NMR data have been included from several transformation products (TPs) and/or parent compounds of contaminants of emerging concern, including: ^{13}C NMR, ^{19}F NMR, ^{1}H NMR, ^{31}P NMR, MS/MS (all from S74 REFTPS [71] containing MS/MS data extracted from 4 articles [258–262] and NMR data from 1 article [258]). Many of these CIDs were not available in PubChem previously. Measured CCS values (often for multiple adducts) associated with 1579 CIDs are included in PubChem, from the datasets S50 CCSCOMPEND [121], S61 UJICCSLIB [150] and S79 UACCSCEC [187] (see also inset at the bottom left in Fig. 5). As mentioned above, this data can be retrieved from PubChem, with documentation provided on the ECI GitLab pages [252–254], along with an archive of the CCS data on Zenodo [255].

Discussion

NORMAN-SLE coverage

The NORMAN-SLE (https://www.norman-network.com/nds/SLE/) provides users with simple access to suspect lists. These lists are then integrated into the merged NORMAN SusDat collection (https://www.norman-network.com/nds/susdat) in the so-called "MS-ready" [263] form (ready for mass spectral screening, i.e., desalted, neutralized, etc.) with a searchable summary table containing NORMAN-relevant additional properties such as fragmentation information, retention time indices [238] and predicted toxicity values [264]. Over the seven years since the launch of the NORMAN-SLE, the website has grown from hosting a handful of lists to now hosting 99 formal referenced collections, amounting to information on 117,071 substances and 115,248 unique compounds (see Table 2). While these total numbers represent only 0.1% of PubChem contents, it is approximately 12% the size of CompTox, i.e., a significant portion of openly available data on environmentally relevant chemicals. Approximately 43,300 CIDs associated with the NORMAN-SLE lists are not yet available in CompTox lists (calculated by overlapping the PubChem NORMAN-SLE and US EPA DSSTox trees on 31 May 2022; documented here [265]). A large proportion of these CIDs missing in CompTox come from the European market lists S32 REACH2017 [86] from the REACH regulation and S17 KEMIMARKET from the Swedish Chemicals Agency (KEMI) [68], as well as from S71 CECSCREEN [85]. It is important to note the discrepancy between the NORMAN-SLE and CompTox versions of NORMAN-SLE lists, especially if the European-relevant chemicals are the focus of suspect screening efforts. This discrepancy results, in part, from the fact that it has been challenging to verify the identities of a large number of the REACH chemicals; many of these are also missing from the PubChemLite collection due to a lack of additional annotation content [241]. Of the 115,248 CIDs integrated in PubChem, 6275 CIDs come exclusively from the NORMAN-SLE (31 May 2022). This highlights that several NORMAN-SLE lists provide valuable data that is not otherwise available in the open domain, including, e.g., mycotoxins that are not commercially available, but have been isolated via fungal fermentation and purification (S26 MYCOTOXINS [200]), as well as newly published PFAS and TPs added via the S46 PFASNTREV19 [46], S74 REFTPS [71] and S96 ECIPFAS [72] lists (among others).

An overview of the number of regulatory lists and major topics is given in Table 4. Key topics include pharmaceuticals, toxins, pesticides, PFAS, TPs, plastics, priority lists, surfactants, and suspect lists for water, with

Table 4 NORMAN-SLE lists (given by suspect list "S" number only for readability) associated with various topics and sources

Topic	Total	Lists	Notes
Regulatory	16	S14, S17, S20, S23, S28, S30, S32, S36, S39, S47, S54, S63, S67, S86, S88, S97	Includes data from ECHA, EFSA, KEMI, REACH, UBA
Pharmaceuticals	11	S6, S10, S44, S45, S55, S56, S57, S58, S72, S76, S92	Excludes personal care products (S13)
Toxins	11	S24^, S26*, S29*, S37^, S40*, S43^, S52, S58, S64*, S75*, S85*	Includes *natural toxins, ^neurotoxins and others
PFAS	10	S9, S14, S25, S46, S80, S89, S92, S94, S95, S96	
TPs	10	S8, S11, S38, S66, S68, S73, S74, S78, S79, S81	
Water	9	S2, S5, S36*, S39, S41, S63, S64, S82*, S84*	Includes *PMT lists
Pesticides	7	S11, S28, S59, S60, S69, S88, S94	
Nat. products	7	S26, S29, S40, S59, S64, S75, S85	
Plastics	6	S20, S47, S48, S49, S67, S97	Plastics/chemicals associated with plastics
Priority	5	S15, S16, S36, S54, S63	Priority monitoring lists
Surfactants	4	S7, S8, S18, S23	

PMT: persistent, mobile, toxic substances. Nat. products: natural products. Other abbreviations: see abbreviation listing

16 lists coming from European regulatory authorities. Future topics are discussed below.

Recognition, role and use of the NORMAN-SLE

The collection of download and view statistics on Zenodo, along with citation tracking, has helped track the impact of the NORMAN-SLE on the community, as shown in Tables 2 and 3. Since the Zenodo integration only commenced in 2019, these statistics only cover a fraction of the real-world use of the NORMAN-SLE. Several efforts known to the authors that build on NORMAN-SLE information are not captured within these statistics, including for instance CECSCREEN [84], which retrieved much of the NORMAN-SLE data that was integrated into CECSCREEN via CompTox. While a PubMed query on the NORMAN-SLE and the sub-collections was attempted to discover more citing articles, this did not return sufficiently reliable results for further interpretation (various text queries generated large numbers of false positives without finding true positives); it seems that environmental literature is not sufficiently covered in PubMed. Guidance is now provided on the NORMAN-SLE website to help users correctly cite the works; it is hoped that this publication will also help to raise awareness of the resource for the wider scientific community—and will highlight the necessity to cite contributions, so that the level of community adoption becomes more visible over time.

The unique views, downloads, and citations available on Zenodo revealed some surprising results. While in NORMAN much focus was given to pesticides, pharmaceuticals, REACH registered chemicals and TP lists due to popular demand, the most popular list by far (see Table 3) proved to be S13 EUCOSMETICS [108], a collection of chemicals employed in cosmetics from EU regulations [106, 107]. The second most viewed list was a Swiss pesticide and metabolite list, S60 SWISSPEST19 [129], a quite recent collection by Kiefer et al. [128] from Eawag, which was expected to gain significant attention. This was an updated version of S11 SWISSPEST [189] from Moschet et al. [12]. While the NORMAN-SLE has several pharmaceuticals lists, the third most viewed list—a pharmaceuticals list, S72 NTUPHTW—was in fact a 2021 contribution from the National Taiwan University (Chen et al. [146]), which was received following a peer-review recommendation for submission to the NORMAN-SLE during manuscript revisions. This was the first such external contribution and marks a milestone in the NORMAN-SLE development. While S0 SUSDAT only appeared in 5th place according to views/downloads, these numbers are only a small fraction of the real statistics, since NORMAN SusDat is also available on a dedicated interactive website. This is also reflected in the relatively high citation count for SusDat compared with other lists. The NORMAN SusDat website (https://www.norman-network.com/nds/susdat/) was visited 120,221 times (20,258 times counting unique IP addresses per day) between 27 Feb. 2020 and 13 July 2022, compared with 26,318 visits to the NORMAN-SLE website (https://www.norman-network.com/nds/SLE/). The original versions of two highly popular lists, the Food Contact Chemicals database (FCCdb) and the database of Chemicals associated with Plastic Packaging (CPPdb) are also available on Zenodo. These have much higher views and (for FCCdb only) download statistics associated with their original depositions compared with the NORMAN-SLE version (which directs viewers back to the original resource with a request to cite the original dataset). The numbers (10 July 2022) are (unique views/downloads): CPPdb [103] (2,082/659), S48/S49 CPPDBLISTA/B [104, 152] (594/1041), FCCdb [88] (8,612/3,703), S77 FCCDB [89] (410/398). Neither of these original depositions have

any citations. The reason for the parallel integration of these lists (i.e., an original version plus NORMAN-SLE version) is to ensure the maintenance of the full integration with the NORMAN-SLE website, PubChem and CompTox (as these require the preparation and archive of additional files, as well as the ability to edit the depositions and make any necessary adjustments).

All NORMAN-SLE lists feed into the merged collection NORMAN SusDat, which forms the basis of the NORMAN Database System (NDS) [29, 31] and integration into other NORMAN initiatives such as the Digital Sample Freezing Platform (DSFP) [266] and prioritization efforts (see Fig. 2). Several NORMAN-SLE lists are associated with NORMAN activities such as collaborative trials [4, 6] and NormaNEWS [184, 199]. NORMAN SusDat and the DSFP are used extensively in many studies in Europe (e.g., [142, 237, 240],), many of which are still in the process of being published. Beyond NORMAN activities and the statistics presented above, gauging the impact of the NORMAN-SLE remains rather intangible at present, since much of it also relates to the use of NORMAN SusDat. Anecdotally, the efforts behind the S11 SWISSPEST and S60 SWISSPEST19 lists have led to the inclusion of more compounds in the (Swiss) national monitoring program [267, 268], while the efforts related to S2 STOFFIDENT have resulted in the discovery of new P-containing compounds (unpublished results).

FAIR data and chemical curation

The decision to deposit the NORMAN-SLE collections on Zenodo helped "FAIRify" [22, 23, 269] the NORMAN-SLE via the provision of DOIs and versioning control. This helps trace updates and provide static URLs to data files, enabling powerful automatic integration such as that currently performed with PubChem (see Fig. 2), as well as providing the citation possibilities and statistics presented above. These are all features that are not currently possible via the infrastructure supporting the NORMAN-SLE website. Version control is important to track changes to the lists; not only in terms of fixing errors (i.e., curation), but also to keep historical records of lists as they change, since some chemicals that have, e.g., been phased out in the EU or changed in relevance may still occur in imported products and the environment. Overall, the data in the NORMAN-SLE is currently reasonably FAIR: *Findable* via the DOI and InChIKey for deep indexing; *Accessible* via the download options of Zenodo; *Interoperable* via the use of SMILES and InChI; and *Reusable* via the open license (CC-BY 4.0) and the use of community standards where feasible, exemplified by the PubChem integration. A transition to the standardized templates proposed recently [24, 37, 270] will help FAIRify the NORMAN-SLE further; these templates could also form the basis to help propose a set of chemical identifiers needed to establish unique (chemical) identifiers for the future European Open Data Platform.

While best efforts are made to map NORMAN-SLE contributions to identifiers correctly, the resources are not available for extensive curation efforts such as those performed by CompTox. This is coupled with the current "as is" philosophy, where lists are processed to best represent the data as provided. The versioning offered by Zenodo opens options for quality control and updating of lists, however this is still a very manual process and currently decoupled from updates to NORMAN SusDat—workflow and infrastructure upgrades to resolve this are underway. Since NORMAN-SLE lists are both sourced from and deposited to third party systems, and due to the different release cycles (PubChem updates daily, CompTox approximately annually), different versions of the data result—which can cause confusion. A coherent collaborative and timely process to update and circulate updated lists across the various systems would be beneficial; while this currently works well with the automated updates between PubChem and the NORMAN-SLE, it is not yet possible with CompTox.

As mentioned above, the NORMAN-SLE hosts 99 suspect lists, which are then integrated into the merged NORMAN SusDat collection in the so-called "MS-ready" [263] form (ready for mass spectral screening). Access to "MS-ready" suspect lists [263] is urgently needed to reduce the number of trivial mistakes in suspect screening (such as searching for the exact masses of salts or polymers). However, the fact that many NORMAN-SLE lists contained both the original substances and their MS-ready form caused several problems with the PubChem integration and the subsequent mapping of structures to the expert knowledge contained within the lists (e.g., it is unclear to an automated method which structure is associated with the metadata: the original SMILES, or the MS-ready SMILES form). The integration of NORMAN-SLE content in PubChem and CompTox, along with discussions with developers, contributors and users is helping to develop better solutions to some of the challenges associated with the mapping of various chemical forms over time.

Basic cheminformatics limitations still prevent the complete integration of suspect information, such as dealing with undefined structures for which no InChI or InChIKey exists (e.g., isomeric mixes such as surfactants, where several structures are hidden behind one detected "mass"). Taking examples from biocides, UVCBs of interest include: creosote; reaction products of 5,5-dimethylhydantoin, 5-ethyl-5-methylhydantoin with bromine and chlorine (DCDMH); reaction products of paraformaldehyde and 2-hydroxypropylamine (ratio 1:1); or reaction

products of: glutamic acid and N-(C12-C14-alkyl)propylenediamine (Glucoprotamin). For those examples, mixture indicators or marker compounds associated with the UVCB may help evaluate these compounds. Biocidal polymers include "polyhexamethylene biguanide hydrochloride with a mean number-average molecular weight (Mn) of 1415 and a mean polydispersity (PDI) of 4.7 (PHMB(1415;4.7))" or "Polymer of formaldehyde and acrolein" or "Polymer of NMethylmethanamine (EINECS 204-697-4 with (chloromethyl) oxirane (EINECS 203-439-8)/Polymeric quaternary ammonium chloride (PQ Polymer)", where pyrolysis GC–MS may assist analysis (not yet an explicit focus of the NORMAN-SLE lists). The CompTox team has made some efforts to address cases such as these through the definition of "related structures" and PubChem have released "concepts" to group several compounds related to substances under a given concept name, a topic that will be explored further at the upcoming BioHackathon [271]. The definition of chemical identifiers such as an InChI(Key) describing UVCB substances is highly desirable to ensure that these efforts can be automated. While initial efforts such as the mixture InChI (MInChI) show promise (see e.g., Fig. 3 in [51]), there is room for further developments. Organometallic compounds (e.g., methylmercury compounds, organolead/organotin compounds, cyclic volatile methylsiloxanes, gadolinium compounds used as contrast agents) are cases that can be handled to an extent with the current approaches (although not in an "MS-ready" form). Upcoming InChI developments will hopefully improve the handling of organometallic species in databases in the near future [272]. Further examples related to biocides that are currently beyond the scope of the NORMAN-SLE (but are in part covered by the NDS) include microbial preparations or strains used as biocidal products, where metabarcoding or proteomics (peptide biomarkers) could be used for characterization, along with nanomaterials/nanoplastics.

Future updates: new submissions

As described above, submissions and updates to the NORMAN-SLE were frozen during preparation of this manuscript. In the meantime, both new submissions and expressions of interest to update existing lists have been registered, partially stimulated by reaching out to all contributors during the writing of this work. Updates have been suggested for S17 KEMIMARKET [68], S28 EUBIOCIDES [196] with information from ECHA [273], S34 EXPOSOMEXPL [165, 166] with new data from [274] plus new microbial metabolites [275, 276] and S75 CyanoMetDB [118, 119] (next release due early 2023). Suggestions for new contributions include a list of endocrine disruptors within the activities of PARC, the Proposition 65 (Prop-65) list of chemicals from the California EPA [277], Phenol-Explorer [278–280], the Database on Migrating and Extractable Food Contact Chemicals (FCCmigex) [281], and finally a shale gas suspect list [282] that has been applied in other studies: [283, 284] and will fill a long-identified gap with respect to fracking-related content.

Beyond these new suggested submissions, future developments involve improving the current submission system to the NORMAN-SLE. The current submissions generally rely on personal contacts, with only one submission recommended externally so far (S72 NTUPHTW [147]). Manual work for the NORMAN-SLE team would be reduced if contributors would consider using a template, as described recently [24, 37, 270]. While the evolution of openly available batch services offered by PubChem [40] and CompTox [41] have greatly eased the mapping of contributed lists to include the required information for upload, a further semi-automation of this workflow would ease matters further and is already in planning. However, extensive curation based on CAS as performed by CompTox is currently out of scope of the NORMAN-SLE, which is based on fully open access resources. While a feedback loop between CompTox and the NORMAN-SLE would help the NORMAN-SLE benefit from the CompTox curation, this is not currently possible. A submission system such as that offered by PubChem could be considered in the future, but is currently beyond reach of the resources available for the NORMAN-SLE. While these enhancements would be desirable, overall the current system has held up well for 99 lists so far and more contributions are welcomed by emailing the NORMAN-SLE team as detailed on the website: https://www.norman-network.com/nds/SLE/.

Future updates: potential new features

Beyond the new submissions and processing updates mentioned in the previous section, several new features have been suggested (and are being considered) for the NORMAN-SLE and/or the broader NORMAN Database System. These can be grouped into four major areas reflected in the following paragraphs: experimental, TPs, categorization/use and regulatory.

On the experimental side, additional functionality to account for physical chemical properties such as mass, polarity, likely ionization mode and amenability to either GC or LC would be beneficial, along with the link to available MS/MS data and/or reference standards for further confirmation. This information is included to a large extent in NORMAN SusDat, which provides a centralized access point for this information, along with predicted toxicity values [264] and retention indices [238],

but will be streamlined and automated further, also to account for possibilities arising from the PubChem integration. Documentation on how to obtain some of this information via PubChem is also available, e.g., for MS/MS [252] and CCS values [253–255]. Advanced Entrez queries (via PubChem) can be used to limit this to certain measurement modes. Another suggested enhancement related to UVCBs would be to include important substructures such as the head group of surfactants or repeating unit of polymers, which could be linked to MS/MS fragments.

A large focus has been placed on TPs over the recent years. A continuation of ongoing efforts will include adding more TPs, including the extraction of data from literature to fill data gaps [71, 174, 205] and the integration of workflows in patRoon [257] in a manner compatible with other NTS workflows. Over the years, there has been increasing interest to add lists of predicted TPs to the NORMAN-SLE, with submissions including predicted TPs for S6 ITNANTIBIOTICS [159], S71 CECSCREEN [85] (both generated with BioTransformer [111]) and S38 SOLNSLMCTPS [102]. While such lists are valuable for researchers performing NTS, these can cause problems with downstream integration with the NDS, CompTox and PubChem as these predicted structures are not necessarily observed and verified, while the number of entries can be an order of magnitude higher (or more) than the source list. These datasets are generally decoupled from the cross-integration at present. A future discussion for NORMAN will be how best to integrate predicted TP data, with the possibility of a "Transformations" module to be added—potentially to represent both documented transformations (e.g., similarly as shown in the insets in Figs. 2 and 5) and predicted transformations.

As the NORMAN-SLE list numbers climb, and with several contributions covering related topics (see Table 4), further refinements will be needed to group lists together and allow the selection of certain subsets for different use cases, or the sorting of lists by categories. The extensive integration with PubChem and the resulting need for organization of NORMAN-SLE content in both CompTox and PubChem has given rise to categorization and classification efforts, and preliminary functionality allowing this is already integrated into NORMAN SusDat. Since there is great interest in the gathering of "Use" information and categorization in general, NORMAN has already initiated activities within the Prioritization working group [285] to define and collect relevant use information and categories from members. These activities will feed into subsequent future developments within NORMAN, PARC [28, 29], EU projects such as ZeroPM [229] and beyond.

The NORMAN-SLE is a community resource built on an incredible amount of volunteer effort and rather limited financial resources. The entire NDS is supported through the NORMAN Association and project funding obtained by individual contributors. The integration with external resources such as PubChem, CompTox and Zenodo provides significant added value beyond the capabilities available to NORMAN. This approach is key to foster cooperation among existing regulatory frameworks, helping to share data and improve chemical risk assessment in the shift towards a "one substance, one assessment" paradigm [286]. With the EU strongly supporting Open and FAIR data, including large initiatives such as PARC [28, 29] and EIRENE [30], along with Green Deal projects such as ZeroPM [229], opportunities for further developments, consolidation and harmonization with broader EU efforts, including the future Open Data Platform appear promising. While the idea behind the NORMAN-SLE has broad support, the current infrastructure and personnel could not currently support, for instance, a requirement to host and thus make all European environmental research data Open and FAIR. If, however, the experiences in building the NORMAN-SLE could help contribute towards establishing such a platform (to which the NORMAN-SLE could contribute), this would be a huge benefit for research and researchers.

Conclusions

The NORMAN Suspect List Exchange (NORMAN-SLE) was created to provide a service to NORMAN members and the greater scientific community, in response to a clear need identified in the NORMAN Non-target Collaborative Screening Trial [4]. Through the provision of a centralized website to collect various suspect lists and references, information exchange is ensured to apply the "screen smart" strategy on specific scientific questions. This FAIRified resource is archived on Zenodo to give DOIs for each set, allowing the cross-integration with other resources and formal citation of datasets, raising the profile of the research of various contributors. The combined list formed from all NORMAN-SLE contributions, NORMAN SusDat, serves as a basis for chemical management for the entire NORMAN Database System (NDS), including the NORMAN Digital Sample Freezing Platform (DSFP) [266].

The NORMAN-SLE is not intended to replace major open compound databases such as ChemSpider, PubChem or CompTox, but rather offers a specialized, complementary service targeted to the environmental science community, particularly in relation to suspect screening, for integration within these larger resources, as done with CompTox and PubChem. Raising the awareness about relevant suspect screening lists and the

quality issues surrounding suspect screening is vital for improving the identification of contaminants of emerging concern in the environment, biota, and products, thereby helping to reduce the number of molecular unknowns in mass spectrometry analyses and to facilitate more comprehensive chemicals assessments. The NORMAN-SLE welcomes new submissions of suspect lists within the scope, along with other ideas and feedback, as described on the NORMAN-SLE website (https://www.norman-network.com/nds/SLE/).

Abbreviations

ASNT: Abstract Syntax Notation (ASN.1) Text format; ATC: Anatomical Therapeutic Chemical code; CAS: Chemical Abstract Service; CCS: Collision cross section (ion mobility experiments); CDK: Chemistry Development Kit; CECs: Contaminants of Emerging Concern; CID: PubChem Compound Identifier; CPPdb: Chemicals associated with Plastic Packaging database; CSID: ChemSpider Identifier; CSV: Comma Separated Values; DOI: Digital Object Identifier; DSFP: Digital Sample Freezing Platform; DSSTox: Distributed Structure-Searchable Toxicity (database); DTXSID: Distributed Structure-Searchable Toxicity (DSSTox) substance identifier; EC: European Commission; ECHA: European Chemicals Agency; ECI: Environmental Cheminformatics group, University of Luxembourg; EFSA: European Food Safety Authority; EIRENE: Environmental Exposure Assessment Research Infrastructure; EU: European Union; FAIR: Findable, Accessible, Interoperable, Reusable; FCCdb: Food Contact Chemicals database; FCCmigex: Database on Migrating and Extractable Food Contact Chemicals; GC: Gas chromatography; HRMS: High resolution mass spectrometry; InChI: International Chemical Identifier; InChIKey: Hashed form of the International Chemical Identifier; IP: Internet Protocol; JRC: Joint Research Centre; JSON: JavaScript Object Notation; KEMI: Swedish Chemicals Agency; LC: Liquid chromatography; MInChI: Mixture InChI; MS: Mass spectrometry; MS/MS: Tandem mass spectrometry; NDS: NORMAN Database System; NMR: Nuclear magnetic resonance; NORMAN: Network of reference laboratories, research centres and related organisations for monitoring of emerging environmental substances; NORMAN SusDat: NORMAN Substance Database; NORMAN-SLE: NORMAN Suspect List Exchange; NTS: Non-Target screening; PARC: European Partnership for Chemicals Risk Assessment; PFAS: Per- and polyfluoroalkyl substances; PMT: Persistent, mobile and toxic substances; REACH: Registration, Evaluation, Authorisation and Restriction of Chemicals (EU regulation); SDF: Structure Data Format; SEO: Search Engine Optimization; SID: PubChem Substance Identifier; SMILES: Simplified Molecular-Input Line-Entry System; TPs: Transformation products; UBA: German Environment Agency (Umweltbundesamt); US EPA: United States Environmental Protection Agency; UVCBs: Substances of Unknown or Variable Composition, Complex Reaction Products or Biological Materials; XML: Extensible Markup Language; ZeroPM: Zero Pollution of Persistent, Mobile Substances (EU project).

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12302-022-00680-6.

Additional file 1: Summary of the NORMAN-SLE datasets (CSV format) as of 4 May 2022 [81].

Additional file 2: Overview of the NORMAN-SLE website (DOCX format) as of 30 May 2022 [82].

Additional file 3: Summary of Zenodo view and download statistics, plus citations (CSV format) as of 28 April 2022 [235].

Additional file 4: Summary of Zenodo citations plus DOIs per list (CSV format) as of 1 May 2022 [236].

Additional file 5: Authorship contributions and acknowledgements mapped to NORMAN-SLE lists (XLSX format).

Acknowledgements

The authors wish to acknowledge all contributors to the NORMAN-SLE and to the information behind the NORMAN-SLE who are not otherwise mentioned in this article. All authors thank those who contributed to all the open software and web services used in this study that have underpinned these efforts. We gratefully acknowledge the contributions of those we could no longer contact and/or who made contributions without our explicit knowledge. Specifically, the authors wish to acknowledge Anca Baesu (McGill University, Canada, S74), Barbara Günthardt (formerly Eawag/Agroscope, S29), Jan Oltmanns (Forschungs- und Beratungsinstitut Gefahrstoffe GmbH (FoBiG), Germany) and Rosa Sjerps (Oasen, Netherlands, S5, S27) who were all approached to be authors and preferred to be acknowledged, along with Robert Mistrik (HighChem, Slovakia, S19) who was approached to be authors but did not respond. Further, the authors acknowledge Ton van Leerdam (KWR, Netherlands), Sascha Lege (formerly University of Tübingen, Germany, S1), Graham Peaslee (Notre Dame University, USA, S9), Guangbo Qu and Guibin Jiang (Chinese Academy of Sciences, China, S46), Marie-Léonie Bohlen and Markus Schwarz (FoBIG, Germany, S54), Oliver Licht and Sylvia Escher (Frauenhofer ITEM, Germany, S54), David Fabregat-Safont, Maria Ibáñez and Juan Vincente Sancho (University Jaume I, Spain, S61), Raoul Wolf (Norwegian Geotechnical Institute, Norway, S90), the PFAS Analytical Exchange Steering Group members Alun James, Anna Kärrman, Audun Heggelund, Belén González-Gaya, Duncan Gray, Griet Jacobs, Leendert Vergeynst, Noora Perkola, Robert Carter, Stefan van Leeuwen and Ulrich Borchers (S95 [215]) as well as Ann Richard, Chris Grulke and the DSSTox curation team (US EPA, USA). This information is also given in Additional file 5. Thanks to the internal reviewers for their helpful comments.

Disclaimers

PJH retired from NIH NLM in 2020 and is now an NIH Special Volunteer in Toxicology and Environmental Health Sciences at NCBI. Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors (VN, ReS) alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer / World Health Organization. The views expressed in this manuscript are solely those of the authors and do not represent the policies of the U.S. Environmental Protection Agency or other agencies. Mention of trade names of commercial products should not be interpreted as an endorsement by the U.S. Environmental Protection Agency. This work has been internally reviewed at the US EPA and has been approved for publication.

Author contributions

ELS founded, coordinates and maintains the NORMAN-SLE (including the Zenodo and GitLab integration), supported by HMT and PC. JS (host), LC (IT), NA and NG (webmaster) host the SLE website on the NORMAN Database System and provide technical support. RA, NA and NST coordinate predicted values and SusDat merging. VD, JS, JH, NST, NA, ELS, EEB, ELW, PJH, HPA, SF, JaK, TL, MaSe, PvdO, ZW provide(d) strategic input to NORMAN-SLE developments. HMT, RA, NA, JPA, HPHA, RB, NB, LiB, LuB, WB, AC, WLC, PC, LDA, YDF, VD, SF, PGF, AG, BG, JG, KG, SG, PH, PJH, SEHa, FH, EMLJ, TJ, KK, MiK, MaK, JeK, MHL, ML, TL, JL, YL, DML, JonM, ADM, JMcL, ChM, JeM, FM, CaM, JaM, MM, MN, VN, KN, HO, JOB, PO, MO, JAP, CP, NR, TR, PR, HR, ReS, SaerS, MaSch, IS, WS, TS, MaSe, KS, HS, RaS, MaSu, KVT, ST, XT, APvW, RCHV, JJV, PvdO, ZW, AJW, DSW, NST, JH, JS, ELS have made contributions to the SLE content as outlined in Additional file 5. AJW helped curate several lists and is responsible for the ongoing registration of lists into the DSSTox database and for the CompTox integration, in coordination with ELS and HMT. JZ, ELS and EEB designed the PubChem/NORMAN-SLE integration, annotation and classification, which was coded and led by JZ, supported by PAT (web services/infrastructure, curation), BAS (deposition, curation), TC (annotation), QL (synonyms/curation) and PC (FAIRifying lists for annotation). ELW and ELS conceptualized the Zenodo deposition. ELS drafted the manuscript, supported by HMT; all authors revised, read and approved the manuscript and submission.

Funding

The NORMAN-SLE project has received funding from the NORMAN Association via its joint proposal of activities. HMT and ELS are supported by the Luxembourg National Research Fund (FNR) for project A18/BM/12341006. ELS, PC, SEH, HPHA, ZW acknowledge funding from the European Union's

Horizon 2020 research and innovation programme under grant agreement No 101036756, project ZeroPM: Zero pollution of persistent, mobile substances. The work of EEB, TC, QL, BAS, PAT, and JZ was supported by the National Center for Biotechnology Information of the National Library of Medicine (NLM), National Institutes of Health (NIH). JOB is the recipient of an NHMRC Emerging Leadership Fellowship (EL1 2009209). KVT and JOB acknowledge the support of the Australian Research Council (DP190102476). The Queensland Alliance for Environmental Health Sciences, The University of Queensland, gratefully acknowledges the financial support of the Queensland Department of Health. NR is supported by a Miguel Servet contract (CP19/00060) from the Instituto de Salud Carlos III, co-financed by the European Union through Fondo Europeo de Desarrollo Regional (FEDER). MM and TR gratefully acknowledge financial support by the German Ministry for Education and Research (BMBF, Bonn) through the project "Persistente mobile organische Chemikalien in der aquatischen Umwelt (PROTECT)" (FKz: 02WRS1495 A/B/E). LiB acknowledges funding through a Research Foundation Flanders (FWO) fellowship (11G1821N). JAP and JMcL acknowledge financial support from the NIH for CCSCompendium (S50 CCSCOMPEND) via grants NIH NIGMS R01GM092218 and NIH NCI 1R03CA222452-01, as well as the Vanderbilt Chemical Biology Interface training program (5T32GM065086-16), plus use of resources of the Center for Innovative Technology (CIT) at Vanderbilt University. TJ was (partly) supported by the Dutch Research Council (NWO), project number 15747. UFZ (TS, MaK, WB) received funding from SOLUTIONS project (European Union's Seventh Framework Programme for research, technological development and demonstration under Grant Agreement No. 603437). TS, MaK, WB, JPA, RCHV, JJV, JeM and MHL acknowledge HBM4EU (European Union's Horizon 2020 research and innovation programme under the grant agreement no. 733032). TS acknowledges funding from NFDI4Chem—Chemistry Consortium in the NFDI (supported by the DFG under project number 441958208). TS, MaK, WB and EMLJ acknowledge NaToxAq (European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie Grant Agreement No. 722493). S36 and S63 (HPHA, SEH, MN, IS) were funded by the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU) Project No. (FKZ) 3716 67 416 0, updates to S36 (HPHA, SEH, MN, IS) by the German Federal Ministry for the Environment, Nature Conservation, Nuclear Safety and Consumer Protection (BMUV) Project No. (FKZ) 3719 65 408 0. MiK acknowledges financial support from the EU Cohesion Funds within the project Monitoring and assessment of water body status (No. 310011A366 Phase III). The work related to S60 and S82 was funded by the Swiss Federal Office for the Environment (FOEN), KK and JH acknowledge the input of Kathrin Fenner's group (Eawag) in compiling transformation products from European pesticides registration dossiers. DSW and YDF were supported by the Canadian Institutes of Health Research and Genome Canada. The work related to S49, S48 and S77 was funded by the MAVA foundation; for S77 also the Valery Foundation (KG, JaM, BG). DML acknowledges National Science Foundation Grant RUI-1306074. YL acknowledges the National Natural Science Foundation of China (Grant No. 22193051 and 21906177), and the Chinese Postdoctoral Science Foundation (Grant No. 2019M650863). WLC acknowledges research project 108C002871 supported by the Environmental Protection Administration, Executive Yuan, R.O.C. Taiwan (Taiwan EPA). JG acknowledges funding from the Swiss Federal Office for the Environment. AJW was funded by the U.S. Environmental Protection Agency. LuB, AC and FH acknowledge the financial support of the Generalitat Valenciana (Research Group of Excellence, Prometeo 2019/040). KN (S89) acknowledges the PhD fellowship through Marie Skłodowska-Curie grant agreement No. 859891 (MSCA-ETN). Exposome-Explorer (S34) was funded by the European Commission projects EXPOsOMICS FP7-KBBE-2012 [308610]; NutriTech FP7-KBBE-2011-5 [289511]; Joint Programming Initiative FOODBALL 2014–17. CP acknowledges grant RYC2020-028901-I funded by MCIN/AEI/1.0.13039/501100011033 and "ESF investing in your future", and August T Larsson Guest Researcher Programme from the Swedish University of Agricultural Sciences. The work of ML, MaSe, SG, TL and WS creating and filling the STOFF-IDENT database (S2) mostly sponsored by the German Federal Ministry of Education and Research within the RiSKWa program (funding codes 02WRS1273 and 02WRS1354). XT acknowledges The National Food Institute, Technical University of Denmark. MaSch acknowledges funding by the RECETOX research infrastructure (the Czech Ministry of Education, Youth and Sports, LM2018121), the CETOCOEN PLUS project (CZ.02.1.01/0.0/0.0/15_003/0000469), and the CETOCOEN EXCELLENCE Teaming 2 project supported by the Czech ministry of Education, Youth and Sports (No CZ.02.1.01/0.0/0.0/17_043/0009632).

Availability of data and materials

All data integrated in the NORMAN Suspect List Exchange are available from the NORMAN-SLE website (https://www.norman-network.com/nds/SLE/) and on the Zenodo NORMAN-SLE community website (https://zenodo.org/communities/norman-sle) or via the individual DOIs (see Table 1). The merged NORMAN SusDat collection is also available (https://www.norman-network.com/nds/susdat/). Individual lists can be accessed by their code on CompTox, the collection can be found under this search URL (https://comptox.epa.gov/dashboard/chemical-lists?search=NORMAN) or on the NORMAN-SLE website (https://www.norman-network.com/nds/SLE/). The NORMAN-SLE is available as data source in PubChem (https://pubchem.ncbi.nlm.nih.gov/source/23819) and browsable as a classification tree (https://pubchem.ncbi.nlm.nih.gov/classification/#hid=101). Detailed annotation content is available in several PubChem compound records, with an overview on the Data Source page (https://pubchem.ncbi.nlm.nih.gov/source/23819). The code supporting the NORMAN-SLE including documentation is available on GitLab (https://gitlab.lcsb.uni.lu/eci/NORMAN-SLE/), along with the code supporting the NORMAN-SLE/PubChem integration (https://gitlab.lcsb.uni.lu/eci/pubchem).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, 6 Avenue du Swing, 4367 Belvaux, Luxembourg. [2]Laboratory of Analytical Chemistry, Department of Chemistry, National and Kapodistrian University of Athens, Panepistimiopolis Zografou, 15771 Athens, Greece. [3]Environmental Institute, Okružná 784/42, 972 41 Koš, Slovak Republic. [4]Oniris, INRAE, LABERCA, 44307 Nantes, France. [5]Norwegian Geotechnical Institute (NGI), Ullevål Stadion, P.O. Box 3930, 0806 Oslo, Norway. [6]Department of Chemistry, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway. [7]Queensland Alliance for Environmental Health Sciences (QAEHS), The University of Queensland, Woolloongabba, QLD 4102, Australia. [8]Leidos, Research Triangle Park, NC, USA. [9]Toxicological Centre, University of Antwerp, Antwerp, Belgium. [10]Environmental and Public Health Analytical Chemistry, Research Institute for Pesticides and Water, University Jaume I, Castelló, Spain. [11]National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, 8600 Rockville Pike, Bethesda, MD 20894, USA. [12]UFZ, Helmholtz Centre for Environmental Research, Leipzig, Germany. [13]Institute of Ecology, Evolution and Diversity, Goethe University, Frankfurt Am Main, Germany. [14]Swedish University of Agricultural Sciences (SLU), Uppsala, Sweden. [15]Institute of Food Safety and Health, College of Public Health, National Taiwan University, 17 Xuzhou Rd., Zhongzheng Dist., Taipei, Taiwan. [16]Faculty of Chemical and Food Technology, Institute of Information Engineering, Automation, and Mathematics, Slovak University of Technology in Bratislava (STU), Radlinského 9, 812 37 Bratislava, Slovak Republic. [17]Science for Life Laboratory, Department of Environmental Science, Stockholm University, 10691 Stockholm, Sweden. [18]Corteva Agriscience, Indianapolis, IN, USA. [19]INERIS, National Institute for Environment and Industrial Risks, Verneuil en Halatte, France. [20]Swedish Chemicals Agency (KEMI), P.O. Box 2, 172 13 Sundbyberg, Sweden. [21]Institute of Environmental Assessment and Water Research-Severo Ochoa Excellence Center (IDAEA), Spanish Council of Scientific Research (CSIC), Barcelona, Spain. [22]Food Packaging Forum Foundation, Staffelstrasse 10, 8045 Zurich, Switzerland. [23]Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, 8092 Zurich, Switzerland. [24]Eawag, Swiss Federal Institute for Aquatic Science and Technology, Überlandstrasse 133, 8600 Dübendorf, Switzerland. [25]Thermo Fisher Scientific, Dornierstrasse 4, 82110 Germering, Germany. [26]Department of Chemistry, Chemical Biological Centre (KBC), Umeå University, Linnaeus Väg 6, 901 87 Umeå, Sweden. [27]Department Environment and Health, Amsterdam Institute for Life and Environment, Vrije Universiteit, Amsterdam, The Netherlands. [28]Water Research Institute (WRI), Nábr. Arm. Gen. L. Svobodu 5, 81249 Bratislava, Slovak Republic. [29]German Environment

Agency (UBA), Wörlitzer Platz 1, Dessau-Roßlau, Germany. [30]Bavarian Environment Agency, 86179 Augsburg, Germany. [31]Analytisches Forschungsinstitut Für Non-Target Screening GmbH (AFIN-TS), Am Mittleren Moos 48, 86167 Augsburg, Germany. [32]Mass Spec Interpretation Services, 3612 Hemlock Park Drive, Kingsport, TN 37663, USA. [33]State Key Laboratory of Environmental Chemistry and Ecotoxicology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences (SKLECE, RCEES, CAS), No. 18 Shuangqing Road, Haidian District, Beijing 100086, China. [34]Hope College, Holland, MI 49422, USA. [35]University of California, Berkeley, CA, USA. [36]Agilent Technologies, Inc., 5301 Stevens Creek Blvd, Santa Clara, CA 95051, USA. [37]Department of Chemistry, Center for Innovative Technology, Vanderbilt-Ingram Cancer Center, Vanderbilt Institute of Chemical Biology, Vanderbilt Institute for Integrative Biosystems Research and Education, Vanderbilt University, Nashville, TN 37235, USA. [38]Institute for Risk Assessment Sciences (IRAS), Utrecht University, Utrecht, The Netherlands. [39]University Rovira i Virgili, Tarragona, Spain. [40]Biosfer Teslab, Reus, Spain. [41]Nutrition and Metabolism Branch, International Agency for Research On Cancer (IARC), 150 Cours Albert Thomas, 69372 Lyon Cedex 08, France. [42]RECETOX, Faculty of Science, Masaryk University, Kotlářská 2, Brno, Czech Republic. [43]Institute of Legal Medicine and Core Facility Metabolomics, Medical University of Innsbruck, Muellerstrasse 44, Innsbruck, Austria. [44]Technologies for Water Management and Treatment Research Group, Department of Civil Engineering, University of Granada, Campus de Fuentenueva S/N, 18071 Granada, Spain. [45]Institute of Health Research Pere Virgili, Tarragona, Spain. [46]Agriculture and Agri-Food Canada/Agriculture et Agroalimentaire Canada, 1391 Sandford Street, London, ON N5V 4T3, Canada. [47]NILU, Norwegian Institute for Air Research, Kjeller, Norway. [48]Fraunhofer Institute for Molecular Biology and Applied Ecology (Fraunhofer IME), Schmallenberg, Germany. [49]Van't Hoff Institute for Molecular Sciences, University of Amsterdam, P.O. Box 94157, Amsterdam 1090 GD, The Netherlands. [50]Laboratory for Operation Control and Research, Zweckverband Landeswasserversorgung, Am Spitzigen Berg 1, 89129 Langenau, Germany. [51]Environment Agency, Horizon House, Deanery Road, Bristol BS1 5AH, UK. [52]Chemical Contamination of Marine Ecosystems (CCEM) Unit, Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER), Rue de l'Ile d'Yeu, BP 21105, 44311 Cedex 3, Nantes, France. [53]Section for Environmental Chemistry and Physics, Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg C, Denmark. [54]Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands. [55]Technology and Society Laboratory, Empa-Swiss Federal Laboratories for Materials Science and Technology, Lerchenfeldstrasse 5, 9014 St. Gallen, Switzerland. [56]Computational Chemistry and Cheminformatics Branch (CCCB), Chemical Characterization and Exposure Division (CCED), Center for Computational Toxicology and Exposure (CCTE), United States Environmental Protection Agency, 109 T.W. Alexander Drive, Research Triangle Park, NC 27711, USA. [57]Department of Bioinformatics-BiGCaT, NUTRIM, Maastricht University, Maastricht, The Netherlands. [58]University of Alberta, Edmonton, AB T6G 2G3, Canada.

Received: 27 July 2022 Accepted: 24 September 2022
Published online: 21 October 2022

References

1. Krauss M, Singer H, Hollender J (2010) LC–high resolution MS in environmental analysis: from target screening to the identification of unknowns. Anal Bioanal Chem 397:943–951. https://doi.org/10.1007/s00216-010-3608-9
2. Hollender J, Schymanski EL, Singer HP, Ferguson PL (2017) Nontarget screening with high resolution mass spectrometry in the environment: ready to go? Environ Sci Technol 51:11505–11512. https://doi.org/10.1021/acs.est.7b02184
3. Schymanski EL, Jeon J, Gulde R et al (2014) Identifying small molecules via high resolution mass spectrometry: communicating confidence. Environ Sci Technol 48:2097–2098. https://doi.org/10.1021/es5002105
4. Schymanski EL, Singer HP, Slobodnik J et al (2015) Non-target screening with high-resolution mass spectrometry: critical review using a collaborative trial on water analysis. Anal Bioanal Chem 407:6237–6255. https://doi.org/10.1007/s00216-015-8681-7
5. Dulio V, van Bavel B, Brorström-Lundén E et al (2018) Emerging pollutants in the EU: 10 years of NORMAN in support of environmental policies and regulations. Environ Sci Eur 30:5. https://doi.org/10.1186/s12302-018-0135-3
6. Rostkowski P, Haglund P, Aalizadeh R et al (2019) The strength in numbers: comprehensive characterization of house dust using complementary mass spectrometric techniques. Anal Bioanal Chem 411:1957–1977. https://doi.org/10.1007/s00216-019-01615-6
7. Schulze B, van Herwerden D, Allan I et al (2021) Inter-laboratory mass spectrometry dataset based on passive sampling of drinking water for non-target analysis. Sci Data 8:223. https://doi.org/10.1038/s41597-021-01002-w
8. NORMAN Association (2022) NORMAN Interlaboratory Studies Website. https://www.norman-network.com/?q=interlab-studies. Accessed 8 Jul 2022
9. Pourchet M, Debrauwer L, Klanova J et al (2020) Suspect and non-targeted screening of chemicals of emerging concern for human biomonitoring, environmental health studies and support to risk assessment: from promises to challenges and harmonisation issues. Environ Int 139:105545. https://doi.org/10.1016/j.envint.2020.105545
10. Grashow R, Bessonneau V, Gerona RR et al (2020) Integrating exposure knowledge and serum suspect screening as a new approach to biomonitoring: an application in firefighters and office workers. Environ Sci Technol 54:4344–4355. https://doi.org/10.1021/acs.est.9b04579
11. Brack W, Bakker J, de Deckere E et al (2005) MODELKEY. Models for assessing and forecasting the impact of environmental key pollutants on freshwater and marine ecosystems and biodiversity (5 pp). Env Sci Poll Res Int 12:252–256. https://doi.org/10.1065/espr2005.08.286
12. Moschet C, Piazzoli A, Singer H, Hollender J (2013) Alleviating the reference standard dilemma using a systematic exact mass suspect screening approach with liquid chromatography-high resolution mass spectrometry. Anal Chem 85:10312–10320. https://doi.org/10.1021/ac4021598
13. Singer HP, Wössner AE, McArdell CS, Fenner K (2016) Rapid screening for exposure to "non-target" pharmaceuticals from wastewater effluents by combining HRMS-based suspect screening and exposure modeling. Environ Sci Technol 50:6698–6707. https://doi.org/10.1021/acs.est.5b03332
14. Schymanski EL, Singer HP, Longrée P et al (2014) Strategies to characterize polar organic contamination in wastewater: exploring the capability of high resolution mass spectrometry. Environ Sci Technol 48:1811–1818. https://doi.org/10.1021/es4044374
15. Sjerps RMA, Brunner AM, Fujita Y et al (2021) Clustering and prioritization to design a risk-based monitoring program in groundwater sources for drinking water. Environ Sci Eur 33:32. https://doi.org/10.1186/s12302-021-00470-6
16. Brunner AM, Dingemans MML, Baken KA, van Wezel AP (2019) Prioritizing anthropogenic chemicals in drinking water and sources through combined use of mass spectrometry and ToxCast toxicity data. J Hazard Mater 364:332–338. https://doi.org/10.1016/j.jhazmat.2018.10.044
17. Letzel T, Bayer A, Schulz W et al (2015) LC–MS screening techniques for wastewater analysis and analytical data handling strategies: sartans and their transformation products as an example. Chemosphere 137:198–206. https://doi.org/10.1016/j.chemosphere.2015.06.083
18. Peter Suber (2015) Open Access Overview (definition, introduction). http://legacy.earlham.edu/~peters/fos/overview.htm. Accessed 3 Jul 2021
19. Kim S, Chen J, Cheng T et al (2021) PubChem in 2021: new data content and improved web interfaces. Nucleic Acids Res 49:D1388–D1395. https://doi.org/10.1093/nar/gkaa971
20. Pence HE, Williams A (2010) ChemSpider: an online chemical information resource. J Chem Educ 87:1123–1124. https://doi.org/10.1021/ed100697w
21. Williams AJ, Grulke CM, Edwards J et al (2017) The CompTox chemistry dashboard: a community data resource for environmental chemistry. J Cheminform 9:61. https://doi.org/10.1186/s13321-017-0247-6
22. GO FAIR (2021) FAIR Principles. https://www.go-fair.org/fair-principles/. Accessed 23 Mar 2021
23. Wilkinson MD, Dumontier M, IjJ A et al (2016) Comment: the FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3:1–9. https://doi.org/10.1038/sdata.2016.18

24. Schymanski EL, Bolton EE (2022) FAIR-ifying the exposome journal: templates for chemical structures and transformations. Exposome 2:osab006. https://doi.org/10.1093/exposome/osab006
25. European Chemicals Agency (ECHA) (2022) European Chemicals Agency (ECHA). https://www.echa.europa.eu/. Accessed 10 Jul 2022
26. European Food Safety Authority (EFSA) (2022) European Food Safety Authority (EFSA). https://www.efsa.europa.eu/en. Accessed 10 Jul 2022
27. European Commission (Joint Research Centre) (2022) Information Platform for Chemical Monitoring (IPCHEM). https://ipchem.jrc.ec.europa.eu/. Accessed 10 Jul 2022
28. Anses, European Commission (2022) European Partnership for the Assessment of Risks from Chemicals (PARC) - Anses Website. In: Anses-Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (French Agency for Food, Environmental and Occupational Health & Safety). https://www.anses.fr/en/content/european-partnership-assessment-risks-chemicals-parc. Accessed 29 May 2022
29. Dulio V, Koschorreck J, van Bavel B et al (2020) The NORMAN Association and the European Partnership for Chemicals Risk Assessment (PARC): let's cooperate! Environ Sci Eur 32:100. https://doi.org/10.1186/s12302-020-00375-w
30. Masaryk University (2022) Environmental Exposure Assessment Research Infrastructure (EIRENE). https://www.eirene-ri.eu/. Accessed 10 Jul 2022
31. Slobodnik J, Hollender J, Schulze T et al (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
32. NORMAN Association (2022) NORMAN Suspect List Exchange (NORMAN-SLE) Website. https://www.norman-network.com/nds/SLE/. Accessed 29 Apr 2022
33. Weininger D (1988) SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. J Chem Inf Model 28:31–36. https://doi.org/10.1021/ci00057a005
34. Heller S, McNaught A, Stein S et al (2013) InChI—the worldwide chemical structure identifier standard. J Cheminform 5:7. https://doi.org/10.1186/1758-2946-5-7
35. American Chemical Society (2022) CAS REGISTRY—the CAS substance collection. https://www.cas.org/cas-data/cas-registry. Accessed 2 Feb 2022
36. European Chemicals Agency (ECHA) (2022) EC inventory. https://www.echa.europa.eu/information-on-chemicals/ec-inventory. Accessed 20 Jun 2022
37. Schymanski EL, Bolton EE (2021) FAIR chemical structures in the Journal of Cheminformatics. J Cheminform 13:50. https://doi.org/10.1186/s13321-021-00520-4
38. Stravs MA, Schymanski EL, Singer HP, Hollender J (2013) Automatic recalibration and processing of tandem mass spectra using formula annotation: recalibration and processing of MS/MS spectra. J Mass Spectrom 48:89–99. https://doi.org/10.1002/jms.3131
39. Schymanski E (2022) RChemMass. https://github.com/schymane/RChemMass. Accessed 27 Apr 2022
40. NCBI/NLM/NIH (2022) PubChem Identifier Exchange. https://pubchem.ncbi.nlm.nih.gov/idexchange/idexchange.cgi. Accessed 23 Jul 2022
41. United States Environmental Protection Agency (2022) CompTox Batch Search. https://comptox.epa.gov/dashboard/dsstoxdb/batch_search. Accessed 23 Jul 2022
42. O'Boyle NM, Banck M, James CA et al (2011) Open Babel: an open chemical toolbox. J Cheminform 3:33. https://doi.org/10.1186/1758-2946-3-33
43. Willighagen EL, Mayfield JW, Alvarsson J et al (2017) The Chemistry Development Kit (CDK) v20: atom typing, depiction, molecular formulas, and substructure searching. J Cheminform 9:33. https://doi.org/10.1186/s13321-017-0220-4
44. Ruttkies C, Schymanski EL, Wolf S et al (2016) MetFrag relaunched: incorporating strategies beyond in silico fragmentation. J Cheminform 8:3. https://doi.org/10.1186/s13321-016-0115-9
45. Trier X, Lunderberg D (2015) S9 | PFASTRIER|PFAS Suspect List: fluorinated substances. Zenodo. https://doi.org/10.5281/zenodo.2621989
46. Liu Y, D'Agostino L, Schymanski E, Martin J (2019) S46|PFASNTREV19|List of PFAS reported in Non-Target HRMS Studies (Liu et al 2019). Zenodo. https://doi.org/10.5281/zenodo.2656744
47. Liu Y, D'Agostino LA, Qu G et al (2019) High-resolution mass spectrometry (HRMS) methods for nontarget discovery and characterization of poly- and per-fluoroalkyl Substances (PFASs) in environmental and human samples. TrAC Trends Anal Chem 121:115420. https://doi.org/10.1016/j.trac.2019.02.021
48. Little J (2017) S18 | TSCASURF|TSCA surfactants. Zenodo. https://doi.org/10.5281/zenodo.2628792
49. Gago-Ferrero P, Schymanski EL, Bletsou AA et al (2015) Extended suspect and non-target strategies to characterize emerging polar organic contaminants in raw wastewater with LC-HRMS/MS. Environ Sci Technol 49:12333–12341. https://doi.org/10.1021/acs.est.5b03454
50. Schymanski EL, Williams AJ (2017) Open science for identifying "known unknown" chemicals. Environ Sci Technol 51:5357–5359. https://doi.org/10.1021/acs.est.7b01908
51. Lai A, Clark AM, Escher BI et al (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. https://doi.org/10.1021/acs.est.2c00321
52. European Organization For Nuclear Research, OpenAIRE, CERN (2013) Zenodo. https://www.zenodo.org/. Accessed 23 Jul 2022
53. NORMAN Association (2022) NORMAN Suspect List Exchange: Zenodo Community. https://zenodo.org/communities/norman-sle/. Accessed 23 Jul 2022
54. Arp HPH, Hale SE, Schliebner I, Neumann M (2022) S36|UBAPMT|Prioritised PMT/vPvM substances in the REACH registration database. Zenodo. https://doi.org/10.5281/zenodo.2653212
55. Grulke CM, Williams AJ, Thillanadarajah I, Richard AM (2019) EPA's DSSTox database: history of development of a curated chemistry resource supporting computational toxicology research. Computat Toxicol 12:100096. https://doi.org/10.1016/j.comtox.2019.100096
56. Schymanski EL, Williams AJ (2018) S24|HUMANNEUROTOX|List of Human Neurotoxins. Zenodo. https://doi.org/10.5281/zenodo.2648769
57. Baker NC, Schymanski EL, Williams AJ (2019) S37|LITMINEDNEURO|Neurotoxicants from literature mining PubMed. Zenodo. https://doi.org/10.5281/zenodo.3242298
58. Baker NC, Schymanski EL, Williams AJ (2019) S43|NEUROTOXINS|Neurotoxicants Collection from Public Resources. Zenodo. https://doi.org/10.5281/zenodo.2656729
59. Schymanski EL, Baker NC, Williams AJ et al (2019) Connecting environmental exposure and neurodegeneration using cheminformatics and high resolution mass spectrometry: potential and challenges. Environ Sci Processes Impacts 21:1426–1445. https://doi.org/10.1039/C9EM00068B
60. Wang Z (2018) S25|OECDPFAS|List of PFAS from the OECD. Zenodo. https://doi.org/10.5281/zenodo.2648776
61. OECD (2018) Toward a new comprehensive global database of per- and polyfluoroalkyl substances (PFASs): Summary report on updating the OECD 2007 list of per- and polyfluorinated substances (PFASs). OECD Report ENV/JM/MONO(2018)7:24
62. US EPA, OECD (2020) CompTox Chemicals Dashboard|PFASOECD Chemicals. https://comptox.epa.gov/dashboard/chemical-lists/PFASOECD. Accessed 29 Dec 2021
63. Williams A (2019) S45|SYNTHCANNAB|Synthetic Cannabinoids from CompTox. Zenodo. https://doi.org/10.5281/zenodo.2656740
64. Epa US, Williams A, Schymanski E (2019) S58|PSCYHOCANNAB|NPS and Synthetic Cannabinoids from CompTox. Zenodo. https://doi.org/10.5281/zenodo.3247723
65. Lowe CN, Williams AJ (2021) Enabling high-throughput searches for multiple chemical data using the U.S.-EPA CompTox Chemicals Dashboard. J Chem Inf Model 61:565–570. https://doi.org/10.1021/acs.jcim.0c01273
66. Schymanski EL, Zhang J, Bolton EE (2022) NORMAN-SLE/PubChem Deposition Mapping File. In: ECI GitLab Pages. https://gitlab.lcsb.uni.lu/eci/pubchem/-/blob/master/deposition/NORMAN_SLE_mappings.txt. Accessed 30 Apr 2022
67. NCBI/NLM/NIH (2022) PubChem Documentation. https://pubchemdocs.ncbi.nlm.nih.gov/about. Accessed 1 May 2022

68. Fischer S (2017) S17|KEMIMARKET|KEMI Market List. Zenodo. https://doi.org/10.5281/zenodo.2628787
69. Association NORMAN, Aalizadeh R, Alygizakis N et al (2018) S0|SUSDAT|Merged NORMAN Suspect List: SusDat. Zenodo. https://doi.org/10.5281/zenodo.2664078
70. Schymanski EL, Li Q, Bolton EE (2022) NORMAN-SLE / PubChem Synonym File. In: ECI GitLab Pages. https://gitlab.lcsb.uni.lu/eci/pubchem/-/blob/master/deposition/SLE_Synonyms.csv. Accessed 30 Apr 2022
71. Schymanski E, Baesu A, Chirsir P (2022) S74|REFTPS|Transformation Products and Reactions from Literature. Zenodo. https://doi.org/10.5281/zenodo.4318838
72. Chirsir P, Schymanski E (2022) S96|ECIPFAS|Updatable List to add PFAS Structures to Public Resources from ECI (UniLu). Zenodo. https://doi.org/10.5281/zenodo.6389740
73. NORMAN Association, NCBI/NLM/NIH (2022) NORMAN-SLE Data Source in PubChem. https://pubchem.ncbi.nlm.nih.gov/source/23819. Accessed 23 Jul 2022
74. Zhang J, Schymanski EL, Thiessen PA, Bolton EE (2022) NORMAN Suspect List Exchange Tree on PubChem Classification Browser. https://pubchem.ncbi.nlm.nih.gov/classification/#hid=101. Accessed 30 Apr 2022
75. Schymanski EL, Zhang J, Bolton EE (2022) NORMAN-SLE / PubChem Classification Mapping File. In: ECI GitLab Pages. https://gitlab.lcsb.uni.lu/eci/pubchem/-/blob/master/deposition/NORMAN_SLE_classification.txt. Accessed 30 Apr 2022
76. Schymanski EL, LCSB-ECI, NCBI/NLM/NIH (2022) LCSB-ECI/PubChem Documentation. In: ECI GitLab Pages. https://gitlab.lcsb.uni.lu/eci/pubchem-docs
77. Schymanski EL (2022) Converting NORMAN-SLE lists to SDF via PubChem. In: ECI GitLab Pages. https://gitlab.lcsb.uni.lu/eci/NORMAN-SLE/-/blob/master/docs/SLEtoSDF.md. Accessed 10 Jul 2022
78. NCBI/NLM/NIH (2022) PubChem Table of Contents Classification Browser. https://pubchem.ncbi.nlm.nih.gov/classification/#hid=72. Accessed 23 Jul 2022
79. Kim S, Cheng T, He S et al (2022) PubChem protein, gene, pathway, and taxonomy data collections: bridging biology and chemistry through target-centric views of PubChem data. J Mol Biol 434:167514. https://doi.org/10.1016/j.jmb.2022.167514
80. Schymanski EL, Chirsir P, LCSB-ECI, et al (2022) PubChem Annotation Content. In: ECI GitLab Pages. https://gitlab.lcsb.uni.lu/eci/pubchem/-/tree/master/annotations. Accessed 1 May 2022
81. Schymanski EL (2022) NORMAN-SLE List Overview 2022–05–04 (CSV). In: ECI GitLab Pages. https://gitlab.lcsb.uni.lu/eci/NORMAN-SLE/-/blob/master/stats/NORMAN-SLE_List_Overview_20220504.csv. Accessed 30 May 2022
82. Schymanski EL (2022) NORMAN-SLE Website Overview 2022–05–30 (DOCX). In: ECI GitLab Pages. https://gitlab.lcsb.uni.lu/eci/NORMAN-SLE/-/blob/master/web/NORMAN-SLE_Website_Overview_20220530.docx. Accessed 30 May 2022
83. NORMAN Association (2022) NORMAN Substance Database (NORMAN SusDat) Website. https://www.norman-network.com/nds/susdat/. Accessed 29 Apr 2022
84. Meijer J, Lamoree M, Hamers T et al (2021) An annotation database for chemicals of emerging concern in exposome research. Environ Int 152:106511. https://doi.org/10.1016/j.envint.2021.106511
85. Meijer J, Lamoree M, Hamers T et al (2020) S71|CECSCREEN|HBM4EU CECscreen: screening list for chemicals of emerging concern plus metadata and predicted phase 1 metabolites. Zenodo. https://doi.org/10.5281/zenodo.3956586
86. Alygizakis N, Slobodnik J (2018) S32|REACH2017|>68,600 REACH Chemicals. Zenodo. https://doi.org/10.5281/zenodo.2653021
87. Groh KJ, Geueke B, Martin O et al (2021) Overview of intentionally used food contact chemicals and their hazards. Environ Int 150:106225. https://doi.org/10.1016/j.envint.2020.106225
88. Groh K, Geueke B, Muncke J (2020) FCCdb: food contact chemicals database. Version 5.0. Zenodo. https://doi.org/10.5281/zenodo.4296944
89. Groh K, Geueke B, Chirsir P et al (2021) S77|FCCDB|Food Contact Chemicals Database v5.0. Zenodo. https://doi.org/10.5281/zenodo.4625495
90. Letzel T, Grosse S, Sengel M (2017) S2|STOFFIDENT|HSWT/LfU STOFFIDENT Database of Water-Relevant Substances. Zenodo. https://doi.org/10.5281/zenodo.2621452
91. Mistrik R (2017) S19|MZCLOUD|mzCloud compounds. Zenodo. https://doi.org/10.5281/zenodo.2628861
92. Aalizadeh R (2019) S55|ZINC15PHARMA|>8600 Pharmaceuticals from ZINC15. Zenodo. https://doi.org/10.5281/zenodo.3247749
93. Irwin J (2022) ZINC15. https://zinc15.docking.org/substances/subsets/world-not-fda/. Accessed 29 Apr 2022
94. Sterling T, Irwin JJ (2015) ZINC 15—ligand discovery for everyone. J Chem Inf Model 55:2324–2337. https://doi.org/10.1021/acs.jcim.5b00559
95. Slobodnik J (2018) S33|SOLUTIONSMLOS|Chemicals used for Modelling in SOLUTIONS. Zenodo. https://doi.org/10.5281/zenodo.2653023
96. SOLUTIONS Consortium (2018) Solutions Project Website. https://www.solutions-project.eu/. Accessed 29 Apr 2022
97. Brack W, Altenburger R, Schüürmann G et al (2015) The SOLUTIONS project: challenges and responses for present and future emerging pollutants in land and water resources management. Sci Total Environ 503–504:22–31. https://doi.org/10.1016/j.scitotenv.2014.05.143
98. Sjerps R (2018) S27|KWRSJERPS2|Extended Suspect List from Sjerps et al (KWRSJERPS). Zenodo. https://doi.org/10.5281/zenodo.2648818
99. Sjerps RMA, Vughs D, van Leerdam JA et al (2016) Data-driven prioritization of chemicals for various water types using suspect screening LC-HRMS. Water Res 93:254–264. https://doi.org/10.1016/j.watres.2016.02.034
100. Ng K, Alygizakis N, Androulakakis A et al (2022) Target and suspect screening of 4777 per- and polyfluoroalkyl substances (PFAS) in river water, wastewater, groundwater and biota samples in the Danube River Basin. J Hazard Mater 436:129276. https://doi.org/10.1016/j.jhazmat.2022.129276
101. Ng K, Alygizakis N, Slobodnik J (2021) S89|PRORISKPFAS|List of PFAS Compiled from NORMAN SusDat. Zenodo. https://doi.org/10.5281/zenodo.5769582
102. LMC (Several Project Partners) (2019) S38|SOLNSLMCTPS|SOLUTIONS Predicted Transformation Products by LMC. Zenodo. https://doi.org/10.5281/zenodo.2653560
103. Groh KJ, Backhaus T, Carney-Almroth B et al (2018) Database of chemicals associated with plastic packaging (Cppdb), Updated Oct 9, 2018. Zenodo. https://doi.org/10.5281/zenodo.1287773
104. Groh K, Schymanski E (2019) S49|CPPDBLISTB|Database of Chemicals possibly (List B) associated with Plastic Packaging (CPPdb). Zenodo. https://doi.org/10.5281/zenodo.2658152
105. Groh KJ, Backhaus T, Carney-Almroth B et al (2019) Overview of known plastic packaging-associated chemicals and their hazards. Sci Total Environ 651:3253–3268. https://doi.org/10.1016/j.scitotenv.2018.10.015
106. The Scientific Committee on Cosmetic Products and Non-Food Products Intended for Consumers (SCCNFP) (2000) The 1st Update of the Inventory of Ingredients Employed in Cosmetic Products. SECTION II: Perfume and Aromatic Raw Materials. In: Report SCCNFP/0389/00 Final. https://www.norman-network.com/sites/default/files/files/suspectListExchange/SCCNFP038900_INCI-2000.pdf. Accessed 29 Apr 2022
107. European Commission (2006) COMMISSION DECISION of 9 February 2006 amending Decision 96/335/EC establishing an inventory and a common nomenclature of ingredients employed in cosmetic products (2006/257/EC). Official Journal of the European Union 2006/257/EC:528
108. von der Ohe P, Aalizadeh R (2017) S13|EUCOSMETICS|Combined Inventory of Ingredients Employed in Cosmetic Products (2000) and Revised Inventory (2006). Zenodo. https://doi.org/10.5281/zenodo.2624119
109. Oswald P, Alygizakis N, Oswaldova M, Slobodnik J (2020) S70|EISUSGCEIMS|Environmental Institute GC-EI-MS suspect list. Zenodo. https://doi.org/10.5281/zenodo.3894827
110. Djoumbou-Feunang Y, Schymanski E, Zhang J, Wishart DS (2020) S73|METXBIODB|Metabolite Reaction Database from BioTransformer. Zenodo. https://doi.org/10.5281/zenodo.4056560
111. Djoumbou-Feunang Y, Fiamoncini J, Gil-de-la-Fuente A et al (2019) BioTransformer: a comprehensive computational tool for small molecule metabolism prediction and metabolite identification. J Cheminform 11:2. https://doi.org/10.1186/s13321-018-0324-5
112. Swedish Chemicals Agency (KEMI) (2015) Occurrence and use of highly fluorinated substances and alternatives. Report from a Government Assignment, Kemikalieinspektionen, Stockholm, Sweden Report 7/15
113. Fischer S (2017) S14|KEMIPFAS|PFAS highly fluorinated substances list: KEMI. Zenodo. https://doi.org/10.5281/zenodo.2621525

114. Alygizakis N (2018) S21|UATHTARGETS|University of Athens Target List. Zenodo. https://doi.org/10.5281/zenodo.2632411
115. Alygizakis NA, Besselink H, Paulus GK et al (2019) Characterization of wastewater effluents in the Danube River Basin with chemical screening, in vitro bioassays and antibiotic resistant genes analysis. Environ Int 127:420–429. https://doi.org/10.1016/j.envint.2019.03.060
116. Horai H, Arita M, Kanaya S et al (2010) MassBank: a public repository for sharing mass spectral data for life sciences. J Mass Spectrom 45:703–714. https://doi.org/10.1002/jms.1777
117. Schymanski E, Schulze T, Alygizakis N (2017) S1|MASSBANK|NORMAN Compounds in MassBank. Zenodo. https://doi.org/10.5281/zenodo.2621391
118. Jones MR, Pinto E, Torres MA et al (2021) CyanoMetDB, a comprehensive public database of secondary metabolites from cyanobacteria. Water Res 196:117017. https://doi.org/10.1016/j.watres.2021.117017
119. Jones MR, Pinto E, Torres MA et al (2021) S75|CyanoMetDB|Comprehensive database of secondary metabolites from cyanobacteria. Zenodo. https://doi.org/10.5281/zenodo.4551528
120. Haglund P, Rostkowski P (2019) S35|INDOORCT16|Indoor Environment Substances from 2016 Collaborative Trial. Zenodo. https://doi.org/10.5281/zenodo.2653206
121. Picache J, McLean J (2019) S50|CCSCOMPEND|The Unified Collision Cross Section (CCS) Compendium. Zenodo. https://doi.org/10.5281/zenodo.2658162
122. Picache JA, McLean JA (2018) Collision Cross Section Database. In: Vanderbilt University. https://lab.vanderbilt.edu/mclean-group/collision-cross-section-database/. Accessed 29 Apr 2022
123. Picache JA, Rose BS, Balinski A et al (2019) Collision cross section compendium to annotate and predict multi-omic compound identities. Chem Sci 10:983–993. https://doi.org/10.1039/C8SC04396E
124. Glüge J, Scheringer M, Cousins IT et al (2021) S80|PFASGLUEGE|Overview of PFAS Uses. Zenodo. https://doi.org/10.5281/zenodo.5029173
125. Glüge J, Scheringer M, Cousins IT et al (2020) An overview of the uses of per- and polyfluoroalkyl substances (PFAS). Environ Sci Processes Impacts 22:2345–2373. https://doi.org/10.1039/D0EM00291G
126. Phillips K (2018) S22|EPACONS|US EPA Consumer Product Suspect List. Zenodo. https://doi.org/10.5281/zenodo.2648759
127. Phillips KA, Yau A, Favela KA et al (2018) Suspect screening analysis of chemicals in consumer products. Environ Sci Technol 52:3125–3135. https://doi.org/10.1021/acs.est.7b04781
128. Kiefer K, Müller A, Singer H, Hollender J (2019) New relevant pesticide transformation products in groundwater detected using target and suspect screening for agricultural and urban micropollutants with LC-HRMS. Water Res 165:114972. https://doi.org/10.1016/j.watres.2019.114972
129. Kiefer K, Müller A, Singer H, Hollender J (2020) S60|SWISSPEST19|Swiss Pesticides and Metabolites from Kiefer et al 2019. Zenodo. https://doi.org/10.5281/zenodo.3544759
130. Schymanski E (2016) S3|NORMANCT15|NORMAN Collaborative Trial Targets and Suspects. Zenodo. https://doi.org/10.5281/zenodo.2621479
131. Günthardt BF, Hollender J, Hungerbühler K et al (2018) Comprehensive toxic plants-phytotoxins database and its application in assessing aquatic micropollution potential. J Agric Food Chem 66:7577–7588. https://doi.org/10.1021/acs.jafc.8b01639
132. Günthardt B (2018) S29|PHYTOTOXINS|Toxic Plant Phytotoxin (TPPT) Database. Zenodo. https://doi.org/10.5281/zenodo.2652994
133. Postigo C, Gil-Solsona R, Herrera-Batista MF et al (2021) A step forward in the detection of byproducts of anthropogenic organic micropollutants in chlorinated water. Trends Environ Anal Chem 32:e00148. https://doi.org/10.1016/j.teac.2021.e00148
134. Postigo C, Gil-Solsona R, Herrera-Batista MF et al (2021) S87|CHLORINETPS|List of chlorination byproducts of 137 CECs and small disinfection byproducts. Zenodo. https://doi.org/10.5281/zenodo.5767356
135. Oberacher HM (2022) WRTMD or MSforID: Tandem mass spectral identification of small molecules. https://msforid.com/. Accessed 29 Apr 2022
136. Oberacher H (2019) S31|WRTMSD|Wiley Registry of Tandem Mass Spectral Data, MSforID. Zenodo. https://doi.org/10.5281/zenodo.2653017
137. Neuwald I, Muschket M, Zahn D et al (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. https://doi.org/10.1016/j.watres.2021.117645
138. Neuwald I, Muschket M, Zahn D et al (2021) A suspect screening list of 1310 persistent and mobile (PM) candidates. Zenodo. https://doi.org/10.5281/zenodo.5503379
139. Neuwald I, Muschket M, Zahn D et al (2021) S84|UFZHSFPMT|PMT Suspect List from UFZ and HSF. Zenodo. https://doi.org/10.5281/zenodo.5535287
140. Dulio V, Aalizadeh R (2017) S16|FRENCHLIST|French Monitoring List. Zenodo. https://doi.org/10.5281/zenodo.2624325
141. Krauss M, Schulze T (2019) S53|UFZWANATARG|Target Compounds from UFZ WANA. Zenodo. https://doi.org/10.5281/zenodo.3365549
142. Kiefer K, Du L, Singer H, Hollender J (2021) Identification of LC-HRMS nontarget signals in groundwater after source related prioritization. Water Res 196:116994. https://doi.org/10.1016/j.watres.2021.116994
143. Kiefer K, Du L, Singer H, Hollender J (2021) S82|EAWAGPMT|PMT Suspect List from Eawag. Zenodo. https://doi.org/10.5281/zenodo.5500131
144. Alygizakis N (2018) S23|EIUBASURF|Surfactant Suspect List from EI and UBA. Zenodo. https://doi.org/10.5281/zenodo.2648765
145. Fischer S (2019) S39|KEMIWWSUS|Wastewater Suspect List based on Swedish Product Data. Zenodo. https://doi.org/10.5281/zenodo.2653566
146. Chen W-L, Lin S-C, Huang C-H et al (2021) Wide-scope screening for pharmaceutically active substances in a leafy vegetable cultivated under biogas slurry irrigation. Sci Total Environ 750:141519. https://doi.org/10.1016/j.scitotenv.2020.141519
147. Chen W-L (2020) S72|NTUPHTW|Pharmaceutically Active Substances Suspect List from National Taiwan University. Zenodo. https://doi.org/10.5281/zenodo.3955664
148. Wössner A, Singer H (2017) S10|SWISSPHARMA|Pharmaceutical List with Consumption Data. Zenodo. https://doi.org/10.5281/zenodo.2623485
149. Celma A, Sancho JV, Schymanski EL et al (2020) Improving target and suspect screening high-resolution mass spectrometry workflows in environmental analysis by ion mobility separation. Environ Sci Technol 54:15120–15131. https://doi.org/10.1021/acs.est.0c05713
150. Celma A, Fabregat-Safont D, Ibàñez M et al (2019) S61|UJICCSLIB|Collision Cross Section (CCS) Library from UJI. Zenodo. https://doi.org/10.5281/zenodo.3549476
151. Dulio V (2017) S15|NORMANPRI|NORMAN Priority List. Zenodo. https://doi.org/10.5281/zenodo.2624273
152. Groh K, Schymanski E (2019) S48|CPPDBLISTA|Database of Chemicals likely (List A) associated with Plastic Packaging (CPPdb). Zenodo. https://doi.org/10.5281/zenodo.2658143
153. Kirchner M, Alygizakis N (2019) S51|WRIGCHRMS|GC-HRMS target list of WRI. Zenodo. https://doi.org/10.5281/zenodo.2658169
154. Singh RR, Lai A, Krier J et al (2021) Occurrence and distribution of pharmaceuticals and their transformation products in Luxembourgish surface waters. ACS Environ Au 1:58–70. https://doi.org/10.1021/acsenvironau.1c00008
155. Singh RR (2021) S76|LUXPHARMA|Pharmaceuticals Marketed in Luxembourg. Zenodo. https://doi.org/10.5281/zenodo.4587355
156. Ruttkies C, Schymanski EL, Strehmel N et al (2019) Supporting non-target identification by adding hydrogen deuterium exchange MS/MS capabilities to MetFrag. Anal Bioanal Chem 411:4683–4700. https://doi.org/10.1007/s00216-019-01885-0
157. Schymanski E, Krauss M (2019) S42|HDXNOEX|Hydrogen Deuterium Exchange (HDX) Standard Set. Zenodo. https://doi.org/10.5281/zenodo.2656724
158. Paulus GK, Hornstra LM, Alygizakis N et al (2019) The impact of on-site hospital wastewater treatment on the downstream communal wastewater system in terms of antibiotics and antibiotic resistance genes. Int J Hyg Environ Health 222:635–644. https://doi.org/10.1016/j.ijheh.2019.01.004
159. Alygizakis N (2016) S6|ITNANTIBIOTIC|Antibiotic List: ITN MSCA ANSWER. Zenodo. https://doi.org/10.5281/zenodo.2621957
160. Bade R, Bijlsma L, Miller TH et al (2015) Suspect screening of large numbers of emerging contaminants in environmental waters using artificial neural networks for chromatographic retention time

prediction and high resolution mass spectrometry data analysis. Sci Total Environ 538:934–941. https://doi.org/10.1016/j.scitotenv.2015.08.078
161. Bade R, Schymanski E (2015) S4|UJIBADE|University of Jaume I Bade et al List. Zenodo. https://doi.org/10.5281/zenodo.2621917
162. Schollée JE, Schymanski EL, Stravs MA et al (2017) Similarity of high-resolution tandem mass spectrometry spectra of structurally related micropollutants and transformation products. J Am Soc Mass Spectrom 28:2692–2704. https://doi.org/10.1007/s13361-017-1797-6
163. Schollee J, Schymanski E (2020) S66|EAWAGTPS|Parent-Transformation Product Pairs from Eawag. Zenodo. https://doi.org/10.5281/zenodo.3754448
164. International Agency for Research on Cancer (IARC) (2022) Exposome-Explorer: database on biomarkers of environmental exposures. http://exposome-explorer.iarc.fr/. Accessed 29 Apr 2022
165. Neveu V, Salek R, Williams AJ, Schymanski EL (2019) S34|EXPOSOMEXPL|Biomarkers from Exposome-Explorer. Zenodo. https://doi.org/10.5281/zenodo.2653032
166. Neveu V, Moussy A, Rouaix H et al (2017) Exposome-Explorer: a manually-curated database on biomarkers of exposure to dietary and environmental factors. Nucleic Acids Res 45:D979–D984. https://doi.org/10.1093/nar/gkw980
167. Ogawa Y, Tokunaga E, Kobayashi O et al (2020) Current contributions of organofluorine compounds to the agrochemical industry. iScience 23:101467. https://doi.org/10.1016/j.isci.2020.101467
168. Ogawa Y, Tokunaga E, Kobayashi O et al (2022) S94|FLUOROPEST|List of 423 FRAC/HRAC/IRAC classified fluoro-agrochemicals. Zenodo. https://doi.org/10.5281/zenodo.6201559
169. European Chemicals Agency (ECHA) (2022) Mapping exercise—Plastic additives initiative—ECHA. https://echa.europa.eu/mapping-exercise-plastic-additives-initiative. Accessed 29 Apr 2022
170. ECHA (2019) S47|ECHAPLASTICS|A list from the plastic additives initiative mapping exercise by ECHA. Zenodo. https://doi.org/10.5281/zenodo.2658139
171. Schymanski E (2014) S7|EAWAGSURF|Eawag Surfactants Suspect List. Zenodo. https://doi.org/10.5281/zenodo.2621972
172. Menger F, Boström G, Jonsson O et al (2021) Identification of pesticide transformation products in surface water using suspect screening combined with national monitoring data. Environ Sci Technol 55:10343–10353. https://doi.org/10.1021/acs.est.1c00466
173. Menger F, Boström G (2021) S78|SLUPESTTPS|Pesticides and TPs from SLU, Sweden. Zenodo. https://doi.org/10.5281/zenodo.4687924
174. Krier J, Singh RR, Kondić T et al (2022) Discovering pesticides and their TPs in Luxembourg waters using open cheminformatics approaches. Environ Int 158:106885. https://doi.org/10.1016/j.envint.2021.106885
175. Krier J (2020) S69|LUXPEST|Pesticide Screening List for Luxembourg. Zenodo. https://doi.org/10.5281/zenodo.3862688
176. Arp HPH, Hale SE (2019) REACH: Improvement of guidance and methods for the identification and assessment of PMT/vPvM substances. German Environment Agency (UBA) Texte 126/2019:131. ISBN: 1862-4804, Dessau-Roßlau, Germany.
177. Arp HPH, Hale SE, Schliebner I, Neumann M (2022) Prioritised PMT/vPvM substances in the REACH registration database. German Environment Agency (UBA) Texte XXX/2022:(accepted). ISBN: 1862-4804, Dessau-Roßlau, Germany
178. Gago Ferrero P (2016) S8|ATHENSSUS|University of Athens Surfactants and Suspects List. Zenodo. https://doi.org/10.5281/zenodo.2621980
179. Inoue M, Sumii Y, Shibata N (2020) Contribution of organofluorine compounds to pharmaceuticals. ACS Omega 5:10633–10640. https://doi.org/10.1021/acsomega.0c00830
180. Inoue M, Sumii Y, Shibata N (2022) S92|FLUOROPHARMA|List of 340 ATC classified fluoro-pharmaceuticals. Zenodo. https://doi.org/10.5281/zenodo.5979647
181. Trace Analysis and Mass Spectrometry Group (2022) TrAMS: trace analysis and mass spectrometry group. http://trams.chem.uoa.gr/. Accessed 29 Apr 2022
182. Damalas DE, Kokolakis S, Karagiannidis A et al (2020) S65|UATHTARGETSGC|University of Athens GC-APCI-HRMS Target List. Zenodo. https://doi.org/10.5281/zenodo.3753371
183. Alygizakis N, Choi P, Gomez Ramos MJ et al (2020) S62|NORMANEWS2|NormaNEWS2: retrospective screening of new emerging contaminants. Zenodo. https://doi.org/10.5281/zenodo.3634963
184. NORMAN Association (2022) NormaNEWS2 Website. https://www.norman-network.net/?q=node/327. Accessed 29 Apr 2022
185. Mohammed Taha H, Janssen EM-L (2021) S85|MICROCYSTINS|Microcystins from CyanoMetDB. Zenodo. https://doi.org/10.5281/zenodo.5665355
186. Belova L, Caballero-Casero N, van Nuijs ALN, Covaci A (2021) Ion mobility-high-resolution mass spectrometry (IM-HRMS) for the analysis of contaminants of emerging concern (CECs): database compilation and application to urine samples. Anal Chem 93:6428–6436. https://doi.org/10.1021/acs.analchem.1c00142
187. Belova L, Caballero-Casero N, van Nuijs ALN, Covaci A (2021) S79|UACCSCEC|Collision Cross Section (CCS) Library from UAntwerp. Zenodo. https://doi.org/10.5281/zenodo.4704648
188. Galani K, Aligizakis N, Thomaidis N (2019) S57|GREEKPHARMA|Suspect Pharmaceuticals from the National Organization of Medicine, Greece. Zenodo. https://doi.org/10.5281/zenodo.3248883
189. Moschet C (2017) S11|SWISSPEST|Swiss Insecticides, Fungicides and TPs. Zenodo. https://doi.org/10.5281/zenodo.2623741
190. Oltmanns J, Bohlen M, Escher S et al (2019) Final Report: Applying a tested procedure for the identification of potential emerging chemical risks in the food chain to the substances registered under REACH–REACH 2. EFSA Support Publ 16:263. https://doi.org/10.2903/sp.efsa.2019.EN-1597
191. Oltmanns J, Aligizakis N, EFSA, Koschorreck J (2019) S54|EFSAPRI|European Food Safety Authority Priority Substances. Zenodo. https://doi.org/10.5281/zenodo.3248993
192. Fischer S, Rostkowski P (2019) S30|PHENANTIOX|A list of Phenolic Antioxidants from KEMI and NILU. Zenodo. https://doi.org/10.5281/zenodo.2653012
193. Thomaidis NS, Gago-Ferrero P, Ort C et al (2016) Reflection of socioeconomic changes in wastewater: licit and illicit drug use patterns. Environ Sci Technol 50:10065–10072. https://doi.org/10.1021/acs.est.6b02417
194. Alygizakis NA, Gago-Ferrero P, Borova VL et al (2016) Occurrence and spatial distribution of 158 pharmaceuticals, drugs of abuse and related metabolites in offshore seawater. Sci Total Environ 541:1097–1105. https://doi.org/10.1016/j.scitotenv.2015.09.145
195. Alygizakis N, Thomaidis N (2019) S56|UOATARGPHARMA|Target Pharmaceutical/Drug List from University of Athens. Zenodo. https://doi.org/10.5281/zenodo.3248837
196. Rüdel H (2018) S28|EUBIOCIDES|Biocides from the NORMAN Priority List. Zenodo. https://doi.org/10.5281/zenodo.2648820
197. Sjerps R (2016) S5|KWRSJERPS|KWR drinking water suspect list. Zenodo. https://doi.org/10.5281/zenodo.2621942
198. Alygizakis N, Samanipour S, Thomas K (2017) S12|NORMANEWS|NormaNEWS for retrospective screening of new emerging contaminants. Zenodo. https://doi.org/10.5281/zenodo.2623816
199. Alygizakis NA, Samanipour S, Hollender J et al (2018) Exploring the potential of a global emerging contaminant early warning network through the use of retrospective suspect screening with high-resolution mass spectrometry. Environ Sci Technol 52:5135–5144. https://doi.org/10.1021/acs.est.8b00365
200. Renaud J, Sumarah M (2018) S26|MYCOTOXINS|List of Mycotoxins from AAFC. Zenodo. https://doi.org/10.5281/zenodo.2648816
201. Rasmussen A (2016) NaToxAq Project Website. https://natoxaq.ku.dk/. Accessed 29 Apr 2022
202. Schulze T (2020) S64|NATOXAQ|NaToxAq: natural toxins and drinking water quality—from source to tap. Zenodo. https://doi.org/10.5281/zenodo.3695174
203. Aurisano N, Huang L, Milàì Canals L et al (2021) Chemicals of concern in plastic toys. Environ Int 146:106194. https://doi.org/10.1016/j.envint.2020.106194
204. Aurisano N, Huang L, Canals LMI et al (2022) S91| CECTOYS|Chemicals of Emerging Concern (CECs) in plastic toys. Zenodo. https://doi.org/10.5281/zenodo.5933614

205. LCSB-ECI, Krier J, Schymanski E et al (2020) S68|HSDBTPS|Transformation Products Extracted from HSDB Content in PubChem. Zenodo. https://doi.org/10.5281/zenodo.3827487
206. European Commission (2020) COMMISSION REGULATION (EU) 2020/2081 of 14 December 2020 amending Annex XVII to Regulation (EC) No 1907/2006 of the European Parliament and of the Council concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) as regards substances in tattoo inks or permanent make-up. European Commission Regulation C/2020/8758:12
207. European Commission (2008) Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. European Commission Regulation 1272/2008:1355
208. Commission E, Mohammed Taha H, Schymanski E (2021) S86|TATTOOINK|TATTOOINK as per EU regulation 2020/2081. Zenodo. https://doi.org/10.5281/zenodo.5710243
209. US EPA (2022) Chemical Contaminants—CCL 4. https://www.epa.gov/ccl/chemical-contaminants-ccl-4. Accessed 29 Apr 2022
210. Epa US, Schymanski EL, Williams AJ (2019) S41|CCL4|CCL 4 Chemical Candidate List. Zenodo. https://doi.org/10.5281/zenodo.2656716
211. US EPA (2022) Contaminant Candidate List 5 (CCL 5). https://www.epa.gov/ccl/contaminant-candidate-list-5-ccl-5. Accessed 29 Apr 2022
212. Epa US, Schymanski E (2021) S83|CCL5|Contaminant Candidate List CCL 5 (Draft). Zenodo. https://doi.org/10.5281/zenodo.5533801
213. Torres S, Schymanski E, Ramirez N (2019) S52|THSMOKE|Thirdhand Smoke (THS) Compounds. Zenodo. https://doi.org/10.5281/zenodo.2669466
214. Sims K, James A, Kärrman A et al (2022) S95|PFASANEXCH|PFAS List from the NORMAN PFAS Analytical Exchange Activity. Zenodo. https://doi.org/10.5281/zenodo.6384309
215. NORMAN Association, UK Environment Agency, Sims K, PFAS Analytical Exchange Steering Committee (2022) 2021 NORMAN network PFAS Analytical Exchange Final Report. https://www.norman-network.net/sites/default/files/files/QA-QC%20Issues/2021%20NORMAN%20network%20PFAS%20Analytical%20Exchange%20Final%20Report%2014022022.pdf. Accessed 4 Jul 2022
216. Arp HPH, Hale SE (2020) S63|UBADWGW|REACH Registered Substances Detected in Drinking (DW) or Groundwater (GW). Zenodo. https://doi.org/10.5281/zenodo.3637629
217. Aalizadeh R (2019) S59|NPINESCT|Natural Product Insecticides. Zenodo. https://doi.org/10.5281/zenodo.3544741
218. Fischer S (2020) S67|TBUTYLPHENOLS|List of tert-butyl phenols from KEMI. Zenodo. https://doi.org/10.5281/zenodo.3779848
219. German Environment Agency (UBA) (2022) S97|UBABPAALT|List of Bisphenol A Alternatives from UBA. Zenodo. https://doi.org/10.5281/zenodo.6405325
220. Eilebrecht E, Wenzel A, Teigeler M, et al (2020) Bewertung des endokrinen Potenzials von Bisphenol Alternativstoffen in umweltrelevanten Verwendungen (in German): Evaluation of the Endocrine Potential of Bisphenol Alternatives in Environmentally-relevant Uses. German Environment Agency (UBA) Texte 123/2019, Dessau-Roßlau, Germany:88
221. German Environment Agency (UBA) Division IV 1.2 (Biocides) (2021) Empfehlungslisten für die Untersuchung der Umweltbelastung durch Biozide: Aktualisierung der Stofflisten des Berichts UBA-TEXTE 15/2017 (in German): Recommendations to investigate environmental contamination with biocides: updating the chemical lists from UBA-TEXTE 15/2017. German Environment Agency (UBA) Addendum to Texte 114/2017, Dessau-Roßlau, Germany:27
222. German Environment Agency (UBA) Division IV 1.2 (Biocides) (2017) Are biocide emissions into the environment already at alarming levels? Recommendations of the German Environment Agency (UBA) for an approach to study the impact of biocides on the environment. German Environment Agency (UBA) Texte 114/2017, Dessau-Roßlau, Germany:67
223. German Environment Agency (UBA), Mohammed Taha H (2021) S88|UBABIOCIDES|List of Prioritized Biocides from UBA. Zenodo. https://doi.org/10.5281/zenodo.5767494
224. Epa US (2019) S40|ALGALTOX|Algal toxins list from CompTox. Zenodo. https://doi.org/10.5281/zenodo.2656710
225. Swedish Chemicals Agency (KEMI) (2017) Bisfenoler—en kartläggning och analys (in Swedish). EN: Bisphenols—a mapping and analysis. Kemikalieinspektionen, Stockholm, Sweden Rapport 5/17:177
226. Rostkowski P, Fischer S (2017) S20|BISPHENOLS|Bisphenols. Zenodo. https://doi.org/10.5281/zenodo.2631745
227. Merino C, Vinaixa M, Ramirez N (2021) S81|THSTPS|Thirdhand Smoke Specific Metabolites. Zenodo. https://doi.org/10.5281/zenodo.5394629
228. Schymanski E, Wang Z, Wolf R, Arp HPH (2022) S90|ZEROPMBOX1|ZeroPM Box 1 Substances. Zenodo. https://doi.org/10.5281/zenodo.5854251
229. Norwegian Geotechnical Institute (NGI) Welcome to ZeroPM: Zero Pollution of Persistent, Mobile Substances. https://zeropm.eu/. Accessed 29 Apr 2022
230. Schymanski EL, Williams AJ (2019) S44|STATINS|Statins Collection from Public Resources. Zenodo. https://doi.org/10.5281/zenodo.2656736
231. Schymanski, E. & Hakkinen, P. S98|TIRECHEM|Tire-related chemicals in environment from literature, *Zenodo*, https://doi.org/10.5281/zenodo.6405358 (2022).
232. US Environmental Protection Agency (2022) CompTox Chemicals Dashboard: Chemical Lists Page. https://comptox.epa.gov/dashboard/chemical-lists. Accessed 30 May 2022
233. US EPA, NCBI/NLM/NIH (2022) PubChem Classification Browser: EPA DSSTox Tree (PubChem CompTox Chemicals Dashboard Chemical Lists Tree). https://pubchem.ncbi.nlm.nih.gov/classification/#hid=105. Accessed 30 May 2022
234. Schymanski EL, Mohammed Taha H (2022) NORMAN-SLE Repository. In: ECI GitLab Pages. https://gitlab.lcsb.uni.lu/eci/NORMAN-SLE. Accessed 30 May 2022
235. Schymanski EL (2022) NORMAN-SLE Zenodo Statistics 2022-04-28. In: ECI GitLab Pages. https://gitlab.lcsb.uni.lu/eci/NORMAN-SLE/-/blob/master/stats/NORMAN-SLE_Zenodo_stats_20220428.csv. Accessed 30 May 2022
236. Schymanski EL (2022) NORMAN-SLE Zenodo Citations 2022-05-01. In: ECI GitLab Pages. https://gitlab.lcsb.uni.lu/eci/NORMAN-SLE/-/blob/master/stats/NORMAN-SLE_Zenodo_Citations_20220501.csv. Accessed 30 May 2022
237. Nikolopoulou V, Aalizadeh R, Nika M-C, Thomaidis NS (2022) TrendProbe: time profile analysis of emerging contaminants by LC-HRMS non-target screening and deep learning convolutional neural network. J Hazard Mater 428:128194. https://doi.org/10.1016/j.jhazmat.2021.128194
238. Aalizadeh R, Alygizakis NA, Schymanski EL et al (2021) Development and application of liquid chromatographic retention time indices in HRMS-based suspect and nontarget screening. Anal Chem 93:11601–11611. https://doi.org/10.1021/acs.analchem.1c02348
239. McEachran AD, Balabin I, Cathey T et al (2019) Linking in silico MS/MS spectra with chemistry data to improve identification of unknowns. Sci Data 6:141. https://doi.org/10.1038/s41597-019-0145-z
240. Alygizakis N, Konstantakos V, Bouziotopoulos G et al (2022) A multi-label classifier for predicting the most appropriate instrumental method for the analysis of contaminants of emerging concern. Metabolites 12:199. https://doi.org/10.3390/metabo12030199
241. Schymanski EL, Kondić T, Neumann S et al (2021) Empowering large chemical knowledge bases for exposomics: PubChemLite meets MetFrag. J Cheminform 13:19. https://doi.org/10.1186/s13321-021-00489-0
242. Giné R, Capellades J, Badia JM et al (2021) HERMES: a molecular-formula-oriented method to target the metabolome. Nat Methods 18:1370–1376. https://doi.org/10.1038/s41592-021-01307-z
243. Nandika D, Karlinasari L, Arinana A et al (2021) Chemical components of fungus comb from Indo-Malayan termite *Macrotermes gilvus* hagen mound and its bioactivity against wood-staining fungi. Forests 12:1591. https://doi.org/10.3390/f12111591
244. Dekić MS, Radulović NS, Selimović ES, Boylan F (2021) A series of esters of diastereomeric menthols: comprehensive mass spectral libraries and gas chromatographic data. Food Chem 361:130130. https://doi.org/10.1016/j.foodchem.2021.130130
245. Wang Q, Ruan Y, Jin L et al (2021) Target, nontarget, and suspect screening and temporal trends of per- and polyfluoroalkyl

substances in marine mammals from the South China Sea. Environ Sci Technol 55:1045–1056. https://doi.org/10.1021/acs.est.0c06685
246. Brase RA, Schwab HE, Li L, Spink DC (2022) Elevated levels of per- and polyfluoroalkyl substances (PFAS) in freshwater benthic macroinvertebrates from the Hudson River Watershed. Chemosphere 291:132830. https://doi.org/10.1016/j.chemosphere.2021.132830
247. Yukioka S, Tanaka S, Suzuki Y et al (2021) Data-independent acquisition with ion mobility mass spectrometry for suspect screening of per- and polyfluoroalkyl substances in environmental water samples. J Chromatogr A 1638:461899. https://doi.org/10.1016/j.chroma.2021.461899
248. Le Moigne D, Demay J, Reinhardt A et al (2021) Dynamics of the metabolome of Aliinostoc sp. PMC 882.14 in response to light and temperature variations. Metabolites 11:745. https://doi.org/10.3390/metabo11110745
249. Libin Xu Lab (2022) CCSbase: An integrated interface for CCS database and prediction. https://ccsbase.net/. Accessed 23 Jul 2022
250. Ross DH, Cho JH, Xu L (2020) Breaking down structural diversity for comprehensive prediction of ion-neutral collision cross sections. Anal Chem 92:4548–4557. https://doi.org/10.1021/acs.analchem.9b05772
251. Zhang J, Thiessen PA, Schymanski EL et al (2022) PubChem: Aggregated CCS Classification Tree. https://pubchem.ncbi.nlm.nih.gov/classification/#hid=106. Accessed 1 May 2022
252. Schymanski EL (2022) Finding MS(/MS) Information for NORMAN-SLE lists via PubChem. In: ECI GitLab Pages. https://gitlab.lcsb.uni.lu/eci/NORMAN-SLE/-/blob/master/docs/SLEwithMS.md. Accessed 4 Jul 2022
253. Schymanski EL (2022) Finding CCS Values for NORMAN-SLE lists via PubChem. In: ECI GitLab Pages. https://gitlab.lcsb.uni.lu/eci/NORMAN-SLE/-/blob/master/docs/SLEwithCCS.md. Accessed 4 Jul 2022
254. Schymanski EL (2022) Retrieving CCS. In: ECI GitLab Pages. https://gitlab.lcsb.uni.lu/eci/pubchem/-/blob/master/annotations/CCS/CCS_retrieval/RetrievingCCS.pdf. Accessed 4 Jul 2022
255. Schymanski E, Zhang J, Thiessen P, Bolton E (2022) Experimental CCS values in PubChem. Zenodo. https://doi.org/10.5281/zenodo.6800138
256. Schymanski E, Bolton E, Cheng T et al (2021) Transformations in PubChem—full dataset. Zenodo. https://doi.org/10.5281/zenodo.5644560
257. Helmus R, van de Velde B, Brunner AM et al (2022) PatRoon 2.0: improved non-target analysis workflowsincluding automated transformation product screening. JOSS 7:4029. https://doi.org/10.21105/joss.04029
258. Bugsel B, Bauer R, Herrmann F et al (2022) LC-HRMS screening of per- and polyfluorinated alkyl substances (PFAS) in impregnated paper samples and contaminated soils. Anal Bioanal Chem 414:1217–1225. https://doi.org/10.1007/s00216-021-03463-9
259. Martin JW, Mabury SA, O'Brien PJ (2005) Metabolic products and pathways of fluorotelomer alcohols in isolated rat hepatocytes. Chem Biol Interact 155:165–180. https://doi.org/10.1016/j.cbi.2005.06.007
260. Alhelou R, Seiwert B, Reemtsma T (2019) Hexamethoxymethylmelamine—a precursor of persistent and mobile contaminants in municipal wastewater and the water cycle. Water Res 165:114973. https://doi.org/10.1016/j.watres.2019.114973
261. Baesu A, Audet C, Bayen S (2021) Application of non-target analysis to study the thermal transformation of malachite and leucomalachite green in brook trout and shrimp. Curr Res Food Sci 4:707–715. https://doi.org/10.1016/j.crfs.2021.09.010
262. Baesu A, Audet C, Bayen S (2022) Evaluation of different extractions for the metabolite identification of malachite green in brook trout and shrimp. Food Chem 369:130567. https://doi.org/10.1016/j.foodchem.2021.130567
263. McEachran AD, Mansouri K, Grulke C et al (2018) "MS-Ready" structures for non-targeted high-resolution mass spectrometry screening studies. J Cheminform 10:45. https://doi.org/10.1186/s13321-018-0299-2
264. Aalizadeh R, von der Ohe PC, Thomaidis NS (2017) Prediction of acute toxicity of emerging contaminants on the water flea Daphnia magna by Ant Colony Optimization-Support Vector Machine QSTR models. Environ Sci Processes Impacts 19:438–448. https://doi.org/10.1039/C6EM00679E
265. Schymanski EL (2022) Overlap of NORMAN-SLE and CompTox via PubChem. In: ECI GitLab Pages. https://gitlab.lcsb.uni.lu/eci/NORMAN-SLE/-/blob/master/stats/misc/PubChem_CompTox_SLE.md. Accessed 11 Jul 2022
266. Alygizakis NA, Oswald P, Thomaidis NS et al (2019) NORMAN digital sample freezing platform: a European virtual platform to exchange liquid chromatography high resolution-mass spectrometry data and screen suspects in "digitally frozen" environmental samples. TrAC Trends Anal Chem 115:129–137. https://doi.org/10.1016/j.trac.2019.04.008
267. Federal Office for the Environment (FOEN) (2022) Chlorothalonil metabolites in groundwater. https://www.bafu.admin.ch/bafu/en/home/themen/thema-wasser/wasser--fachinformationen/zustand-der-gewaesser/zustand-des-grundwassers/grundwasser-qualitaet/pflanzenschutzmittel-im-grundwasser/chlorothalonil-metaboliten-im-grundwasser.html. Accessed 20 Jul 2022
268. Kiefer K, Müller A, Singer H et al (2019) Pflanzenschutzmittel-metaboliten im Grundwasser (EN: Pesticide Metabolites in Groundwater). Aqua Gas 99:14–23
269. The FAIRsharing Community, Sansone S-A, McQuilton P et al (2019) FAIRsharing as a community approach to standards, repositories and policies. Nat Biotechnol 37:358–367. https://doi.org/10.1038/s41587-019-0080-8
270. NCBI/NLM/NIH (2021) PubChem Submissions Template Folder. https://ftp.ncbi.nlm.nih.gov/pubchem/Other/Submissions/. Accessed 25 May 2021
271. ELIXIR Europe (2022) Project 26: Shedding the light on unknown chemical substances (BioHackathon Europe 2022). In: GitHub. https://github.com/elixir-europe/biohackathon-projects-2022/tree/main/26. Accessed 11 Jul 2022
272. InChI Trust (2022) Organometallics—InChI Trust. https://www.inchi-trust.org/organometallics/. Accessed 11 Jul 2022
273. European Chemicals Agency (ECHA) (2022) Information on biocides—ECHA. https://echa.europa.eu/information-on-chemicals/biocidal-active-substances. Accessed 6 Jul 2022
274. Neveu V, Nicolas G, Salek RM et al (2019) Exposome-Explorer 2.0: an update incorporating candidate dietary biomarkers and dietary associations with cancer risk. Nucleic Acids Res 48:D908–D912. https://doi.org/10.1093/nar/gkz1009
275. International Agency for Research on Cancer (IARC) (2022) Exposome-Explorer: Microbial metabolites. http://exposome-explorer.iarc.fr/microbial_metabolites. Accessed 10 Jul 2022
276. Neveu V, Nicolas G, Amara A et al (2022) The human microbial exposome: expanding the Exposome-Explorer database with gut microbial metabolites. In Review. https://doi.org/10.21203/rs.3.rs-1754003/v2
277. California Office of Environmental Health Hazard Assessment (OEHHA), California Environmental Protection Agency (2022) Proposition 65 Warnings Website - Your right to know. https://www.p65warnings.ca.gov/node. Accessed 6 Jul 2022
278. Neveu V, Perez-Jimenez J, Vos F et al (2010) Phenol-Explorer: an online comprehensive database on polyphenol contents in foods. Database 2010:bap024–bap024. https://doi.org/10.1093/database/bap024
279. Rothwell JA, Urpi-Sarda M, Boto-Ordonez M et al (2012) Phenol-Explorer 2.0: a major update of the Phenol-Explorer database integrating data on polyphenol metabolism and pharmacokinetics in humans and experimental animals. Database 2012:bas031–bas031. https://doi.org/10.1093/database/bas031
280. Rothwell JA, Perez-Jimenez J, Neveu V et al (2013) Phenol-Explorer 3.0: a major update of the Phenol-Explorer database to incorporate data on the effects of food processing on polyphenol content. Database 2013:bat070–bat070. https://doi.org/10.1093/database/bat070
281. Geueke B, Groh KJ, Maffini MV et al (2022) Systematic evidence on migrating and extractable food contact chemicals: most chemicals detected in food contact materials are not listed for use. Crit Rev Food Sci Nutri 1–11. https://doi.org/10.1080/10408398.2022.2067828
282. Faber A-H, Annevelink M, Gilissen HK et al (2017) How to adapt chemical risk assessment for unconventional hydrocarbon extraction related to the water system. In: de Voogt P (ed) Reviews of environmental contamination and toxicology, vol 246. Springer International Publishing, Cham, pp 1–32
283. Faber A-H, Brunner AM, Dingemans MML et al (2021) Comparing conventional and green fracturing fluids by chemical characterisation and effect-based screening. Sci Total Environ 794:148727. https://doi.org/10.1016/j.scitotenv.2021.148727

284. Faber A-H, Annevelink MPJA, Schot PP et al (2019) Chemical and bioassay assessment of waters related to hydraulic fracturing at a tight gas production site. Sci Total Environ 690:636–646. https://doi.org/10.1016/j.scitotenv.2019.06.354
285. NORMAN Association (2022) NORMAN Working Group 1: Prioritisation Website. https://www.norman-network.com/?q=node/50. Accessed 12 Jul 2022
286. van Dijk J, Gustavsson M, Dekker SC, van Wezel AP (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

Publisher's Note