

ChemExpo Knowledgebase User Guide

Version 0.1

Prepared by:

Sakshi Handa, Katherine A. Phillips, Kenta Baron-Furuyama, Kristin K. Isaacs
Center for Computational Toxicology and Exposure
United States Environmental Protection Agency, Office of Research and Development,
Research Triangle Park, NC

In Conjunction with:

General Dynamics Information Technology (GDIT), Research Triangle Park, NC
Contract CIO-SP3, HHSN316201200013W

Last Modified: 5/10/23

Disclaimer

This resource is a compilation of information sourced from many sites, databases, and sources, including U.S. Federal and state sources and international bodies. The data are not reviewed by USEPA - the user must apply judgment in use of the information. You should consult the original data sources, if possible, to ensure an understanding of the context of the data contained in ChemExpo.

Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government. The views and opinions of the developers of the site expressed herein do not necessarily state or reflect those of the United States Government, and shall not be used for advertising or product endorsement purposes.

With respect to documents available from this server, neither the United States Government nor any of their employees, makes any warranty, express or implied, including the warranties of merchantability and fitness for a particular purpose, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.

This User Guide has been subject to Agency review and has been approved for publication.

Table of contents

Table of contents	3
Chapter 1: Purpose of this User Guide	5
Chapter 2: Introduction	6
2.1 Summary of functionality available in V1.0	6
Chapter 3: ChemExpo Data Overview	8
3.1 Chemical Identifiers	9
3.2 Product Use Categories	9
3.3 Function Categories	10
3.4 Chemical List Presence Keywords	10
3.5 Introduction to Data Documents and the Data Document Page	10
Chapter 4: Search	12
Part I: Explore	14
Chapter 5: Explore Chemicals	15
5.1 Product Data	15
5.2 Additional Chemical Information	17
5.3 Chemical List Presence Data	18
5.4 Function Data	18
Chapter 6: Explore Products or Product Use Categories	20
6.1 Product Information in ChemExpo	20
6.2 Browsing PUCs and their Data	20
6.3 PUC Detail Pages	21
6.4 Finding Products via PUCs	22
6.5 Chemical Data	23
Chapter 7: Explore Chemical Function	24
7.1 Browsing Chemical Functions	24
7.2 Chemical Data for a Functional Category	26
7.3 Product Data for a Functional Category	26
7.4 Interpreting FCs	26
Chapter 8: Explore General Chemical Uses	27
8.1 Browse Chemical List Presence Keywords	27
8.2 Data Documents	28
Chapter 9: Explore Data Documents	29
9.1 Accessing Data Documents	29

9.2 Original Data	29
Part II: Download	31
Specific Data	32
9.1 Chemical Data	32
9.2 Product Use Category Data.....	32
9.3 Function Category Data	32
9.4 Chemical List Presence Data.....	33
9.5 Document Data.....	33
Chapter 10: Bulk Data.....	34
10.1 Product Data.....	34
10.2 Function Category Data.....	34
10.3 General Use Data.....	34
Chapter 11: Glossary.....	36
Chapter 12: Frequently Asked Questions (FAQs).....	45
12.1 Data and Maintenance Issues	45
12.2 Usability Questions	47
References.....	49
Part III: Appendices.....	50
Appendix A: Visualization Page Details.....	50
A.1: Chemical Detail Page Tables.....	50
A.2: PUC Detail Page Tables.....	51
A.3: FC Detail Page Tables	53
A.4: List Presence Keyword Detail Page Tables.....	54
A.5: Data Document Tables.....	55
Appendix B: Download Terminology	58
B.1: Metadata Download Fields	58
B.2: Bulk Data Downloads.....	59
B.3: Details Page Downloads.....	63
Appendix C: A Primer on Safety Data Sheets.....	74

Chapter 1: Purpose of this User Guide

This is a guide for users who wish to explore data available on U.S. EPA's Chemical Exposure Knowledgebase (ChemExpo). It is appropriate for both experienced users and those who may be new to the application. The goal is to simplify a user's experience navigating the ChemExpo application by providing examples and use case scenarios.

Welcome to ChemExpo Knowledgebase

ChemExpo is an interactive web application for exploring chemical data, curated from public documents, relevant to exposure assessment. ChemExpo focuses on data collected by EPA about how chemicals are used in commerce and how they occur in consumer and industrial products. ChemExpo provides tools for exploring and downloading these data, which include consumer product composition, chemical functional use, and general chemical use information. The ChemExpo team actively works to curate these data to specific consumer and occupational product categories, to chemical functional uses, and to substance identifiers (DTXSIDs) used by EPA and the CompTox Chemicals Dashboard.



Getting Started

- Use the Search Bar above to search chemical names and identifiers, product names and categories, or documents for a specific text string.
- Use the [Explore Products](#) tab to interactively explore the ChemExpo Product Use Category (PUC) Hierarchy
- Browse the available [Product Use Categories](#)
- Browse the available [Chemical List Presence Keywords](#)
- Browse the available [Chemical Function Categories](#)

Latest News

[See All News](#)

New Bulk Downloads Added

Sept. 7, 2021, 7:35 p.m.

 Bulk Downloads of List Presence, Functional Use and Composition data are now available on the [Get Data](#) page. 

Figure 1.1: ChemExpo's Home Screen

Chapter 2: Introduction

The Chemical Exposure Knowledgebase (ChemExpo) is a new, publicly available web-based data search and visualization tool for the exposure domain developed by the Center for Computational Toxicology and Exposure (CCTE) within U.S. EPA's Office of Research and Development. ChemExpo will surface data managed and curated by the internal-to-U.S. EPA applications and workflows developed to assist with managing, curating, and ensuring the quality of publicly available exposure-relevant data collected by U.S. EPA. These data, which are primarily obtained from documents published by manufacturers, retailers, governments, and NGOs, describe how chemicals are used in commerce, including in consumer products, and are released by EPA as the Chemical and Products database (CPDat) (Dionisio et al. (2018)). CPDat has been previously described in the literature and its data are available in summary form on the [EPA CompTox Chemicals Dashboard](#) (the "CompTox Dashboard"). CPDat contains information on chemical use curated from public documents, organized by chemical, Product Use Category (PUC) (Isaacs et al. (2020)), Function Category (FC) (Phillips et al. (2017)), and Chemical List Presence Keywords (keywords developed to index general information about how chemicals are used from public chemical lists) (Dionisio et al. (2015), Koval et al. (2022)). ChemExpo also makes use of harmonized U.S. EPA chemical identifiers called Distributed Structure-Searchable Toxicity Database (DSSTox) Substance Identifiers (DTXSIDs) (Grulke et al. (2019)), allowing for direct linking to other chemical information in the CompTox Dashboard.

The new ChemExpo application will display more detailed curated CPDat data than currently available on the CompTox Dashboard, including additional document metadata; product-level chemical ingredient information; raw reported information in addition to curated data; and product category, function category, and chemical summaries. All data within ChemExpo were obtained from publicly available primary sources. This application allows for data exploration by chemical, function, PUC, and keyword to support exposure assessments and other risk-based chemical evaluations.

The following sections of this User's Guide provide an overview of the data available in ChemExpo. It also describes how one would search ChemExpo for specific information or use the application's interface to explore the available CPDat data. The guide is organized around specific questions users may have about how to use the application for different purposes. For the definitions of specific terms used in the sections below, see the ChemExpo Glossary.

Data curation within ChemExpo is ongoing, and sometimes makes use of automated methods (noted where appropriate). For more information on the ChemExpo Knowledgebase or to report anything that you feel may be in error contact ChemExpo Support by Email at: chemexpo.support@epa.gov

2.1 Summary of functionality available in V1.0

The initial release of ChemExpo (v1.0) will provide the following functionality. Details of the data, terminology, search functionality, and download procedures are provided in this User Guide.

- Users will be able to search for data related to Documents, Products, Chemicals, Product Use Categories, Chemical List Presence Keywords, and Function Categories.
- Users will be able to explore available data by Chemical, Product Use Category, Chemical List Presence Keyword, and Function Category.
- Users will be able to download data including:
 - Metadata (e.g., definitions) related to Product Use Categories, Function Categories, Chemical List Presence Keywords
 - Bulk downloads of Composition, Function Category, and Chemical List Presence datasets
 - Data downloads related to specific chemicals, Product Use Categories, Function Categories, and Data Documents

Chapter 3: ChemExpo Data Overview

To support chemical decision-making, EPA's Office of Research and Development (ORD) must identify and characterize relevant exposure pathways - the path of a chemical from source to a receptor. How a chemical is used (e.g., in a consumer, occupational, or industrial context) is critical to determining exposure pathways. Over the last decade, ORD has developed a series of datasets and databases containing information collected from public documents that describe how chemicals are used in commerce, including in consumer products. These data are currently released as the Chemicals and Products Database (CPDat). ChemExpo is a new tool that will provide to the public metadata associated with CPDat data sources and new ways to explore and download the data.

This section describes the main classification systems used in ChemExpo and in CPDat to organize chemical use information extracted from public documents. These include harmonized chemical identifiers (DTXSIDs), Product Use Categories (PUCs), Chemical List Presence Keywords (List Keywords), and Function Categories (FCs). Each of these classification systems have been developed over time by U.S. EPA and/or other groups and are described in the peer-reviewed literature. In the following sections we briefly describe the systems, provide some information about the history of their development, and link to both relevant literature and to the formal definitions (e.g., category definitions) associated with each. Document Types in ChemExpo are also defined in this section. Detailed descriptions of the terminology used herein and in the ChemExpo application can be found in the ChemExpo Glossary.

The relationships in ChemExpo among documents, products, the various classification schemes, and various extracted information are summarized in [Figure 3.1](#). These relationships are described in more detail in the following sections.

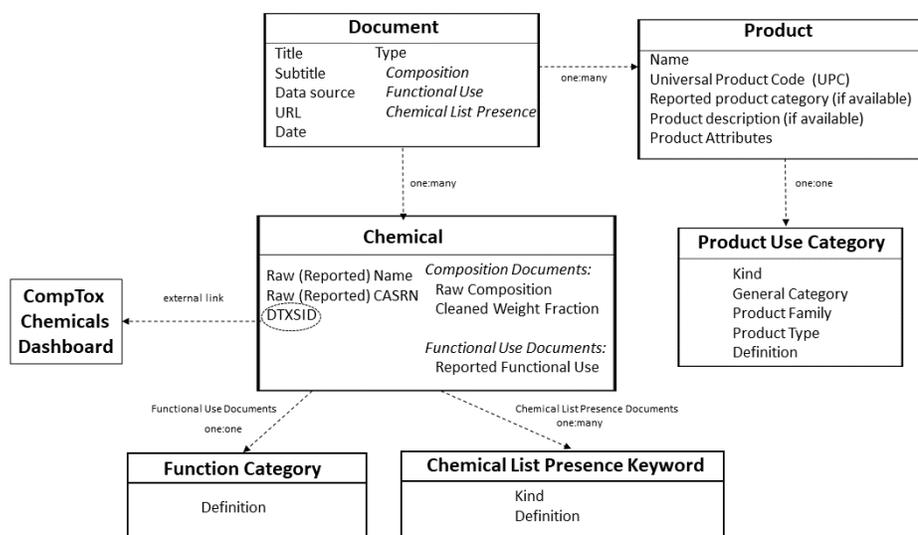


Figure 3.1: ChemExpo's Data Elements

3.1 Chemical Identifiers

ChemExpo makes use of EPA chemical curation workflows (Grulke et al. (2019)) that were developed to support ORD research and the CompTox Dashboard. These workflows match reported chemical identifiers, e.g., chemical names, Chemical Abstracts Service Registry Numbers (CASRNs), etc., to EPA chemical identifiers called Distributed Structure-Searchable Toxicity Database (DSSTox) Substance Identifiers (DTXSIDs) (Grulke et al. (2019)). DTXSIDs can be used to directly link data in ChemExpo to other chemical data (e.g., toxicity information) available in the Dashboard. Chemical curation in ChemExpo is ongoing and in DSSTox; some identifiers may not be able to be successfully curated to DTXSIDs, either because they don't represent a unique substance (e.g., a chemical reported on a document as "fragrance") or because they aren't currently recognized by the curation workflow.

3.2 Product Use Categories

Some data in ChemExpo are curated from documents that describe specific products. These products are organized into Product Use Categories (PUCs) developed explicitly for exposure assessment and modeling. Currently, there are three different kinds of PUC in ChemExpo: those associated with consumer formulations (e.g., cleaners, personal care products), consumer articles (e.g., furniture), and industrial/occupational products (e.g., raw materials, laboratory supplies). The organizational hierarchy for PUCs consists of three levels: General Category, Product Family, and Product Type. PUCs are organized from general to more specific product types, i.e., 'Personal care' is the most general category, while 'Personal care – dental care – toothpaste' would be a specific product type. Products can be curated to a higher level (e.g., General Category or Product Family only) if there is not enough information to assign a product type. The consumer product PUCs used in ChemExpo are those described in Isaacs et al. (2020), with some additions/refinements necessitated by the addition of new products to CPDat. In addition, some very general article and industrial/occupational PUCs have been added to ChemExpo as new data sources for these types of products have been curated. The article PUCs were based on harmonized article categories developed by the Organisation for Economic Co-operation and Development (OECD) (Directorate (2017)) and the industrial/occupational product PUCs were developed ad hoc from examination of the data sources. Products in ChemExpo were curated to PUC manually, in bulk based on data source metadata, or via automated assignment based on product name, using natural language processing methods (peer-reviewed publication in development). In ChemExpo, the type of PUC assignment used (manual, bulk, or automated) is provided for each curated product. Definitions for all PUCs used in ChemExpo are provided on the [PUC Summary Page](#).

Products within PUCs may also be assigned specific Attributes. Attributes are keywords further describing the type of formulation (e.g., liquid, spray), user population (e.g., child), or microenvironment of use (e.g., indoor). These attributes may be of use in an exposure assessment. Some PUCs have assumed attributes (e.g., toys are assumed to be relevant to children), and each PUC has a list of allowable attributes.

3.3 Function Categories

ChemExpo contains data about the technical role, or function, performed by a chemical within products or processes (e.g., solvent, plasticizer, fragrance). These data update those previously included in EPA's Functional Use Database (Phillips et al. (2017)) and CPDat (Dionisio et al. (2018)). In ChemExpo, reported functions are curated to a set of Function Categories (FCs) that include harmonized functional use categories developed by OECD, in addition to other categories added by U.S. EPA for functions not covered by the OECD categories (Directorate (2017)) (EPA-generated FCs are indicated using 'EPA' within the category name). These standardized FCs contain explicit definitions and exclusions, to reduce ambiguity in category assignment. New FCs will be added as needed to accommodate novel uses. Definitions for all FCs used in ChemExpo are on the [FC Summary Page](#).

3.4 Chemical List Presence Keywords

In addition to specific product and function data, ChemExpo contains general information about how chemicals might be used in commerce. These data are indexed by specific keywords (Chemical List Presence Keywords, often abbreviated as "List Keywords" here and in ChemExpo) that define and describe the presence of chemicals on defined lists contained within public documents (e.g., lists of food-use chemicals or chemicals used in specific industries). These keywords are an update/refinement to the terms previously developed for EPA's Chemical and Product Categories (CPCat) database (Dionisio et al. (2015)) and are described in Koval et al. (2022). These refinements better align the assignment of keywords with other CPDat data streams, namely PUCs and FCs. For example, CPCat included specific functions as keywords, as our FC system did not exist at the time. In addition, CPCat contained large number of keywords associated with different consumer product types or categories. These terms have been harmonized and updated to the product use categories (PUCs). For example, a chemical list from a public document, denoting chemicals used in a specific type of personal care product, would be assigned a keyword identical to the PUC that would be assigned to specific products in ChemExpo. Keywords have kinds, as well – there are keywords that modify other keywords and those associated with a geographic location or populations. From a single chemical list, more than one List Keyword can be assigned to a chemical. For example, a list of pesticide active ingredients used in Europe curated from a document would be assigned the keywords "pesticides", "Europe" and "active ingredient". Such a combination is called a List Keyword set. Definitions for all List Keywords (and their associated kind) used in ChemExpo are provided on the [List Presence Keywords Summary Page](#).

3.5 Introduction to Data Documents and the Data Document Page

In ChemExpo, documents are categorized by the type of information they contain. The following document types are used in ChemExpo:

3.5.1 Document Types in ChemExpo

Composition Documents: Composition documents contain lists of chemical ingredients for one or more products, for example, unique Universal Product Codes (UPCs) or product

names. These data may be qualitative (e.g., an ingredient list) or quantitative (e.g., weight fraction information). Multiple product records may be extracted from a single composition document (see [Figure 3.1](#)) and assigned to Product Use Categories, but chemical composition data are associated from the actual document from which it was extracted (since each composition document contains unique composition data, that is, a unique formulation). Composition documents may also contain function information that allow individual chemicals to be mapped to Function Categories (FCs). Composition documents may include Safety Data Sheets (see [Appendix C](#) for more information on these documents), ingredient disclosures, and ingredient lists.

Chemical List Presence Documents: These documents contain one or more chemical lists that are extracted and mapped to one or more Chemical List Presence Keywords (with multiple keywords comprising a keyword set). These reports and documents provide general information on the use of chemicals, and may come from international, federal, or state agencies, trade associations, or other reputable sources.

Functional Use Documents: These documents contain information on functions related to a specific chemical substance. This reported function data allows the chemical to be mapped to FCs. Functional use documents include technical specification sheets, or chemical retailer webpages, or regulatory inventory documents.

3.5.2 The Data Document Page

The Data Document page is the central page type in ChemExpo that contains specifics of the chemical information extracted from a particular document. Data Document pages can be accessed in many ways via Search or via Chemical, Product Use Category, or Function Category pages as described in later sections. The information displayed depends on the type of Data Document. Generally, the page displays document metadata, such as the document title, subtitle, document date, and the reporting organization. Chemical data are organized into 'cards', which contain both raw extracted data and cleaned/harmonized data; these data vary based on Document Type. All chemical cards include the raw (reported) chemical identifiers and DTXSID (if the identifier was successfully curated). On Composition Documents, chemical cards may include raw and cleaned composition information for the chemical. On Functional Use Documents, chemical cards contain raw reported functions for the chemical as well as the curated Function Categories. On Chemical List Presence Documents, chemical cards contain the List Keywords that have been assigned to the chemical based on the definition of the list in the document (see [Table A.5](#) for more information on data available on the Data Document page).

Chapter 4: Search

Tip

Throughout ChemExpo, click on the Info button for more information about a feature or tool.

How do I use the Search functionality in ChemExpo?

ChemExpo uses a single general search box that is located in the main navigation bar, accessible from any page in the application. Users can enter any search query by typing into the search box and pressing the Enter key. ChemExpo searches 6 different categories of information for the entered text: information related to documents, specific products, PUCs, FCs, chemicals, and chemical list keywords. The results (Figure 4.1) are returned on different tabs (one for each type of information). Users can navigate through each of the Search Tabs to view the results. Each tab searches on different data fields, as described in the Table 4.1.

The screenshot displays the ChemExpo search results for the query 'plastic'. At the top, there are six tabs: Documents (919), Products (1362), Product Use Categories (1), Chemicals (54), Chemical List Keywords (2), and Function Categories (4). The search results are displayed in a list format, with each result showing the raw chemical name and the number of products related to the document. A filter sidebar is visible on the left, showing options for Group type: Composition (874), Chemical presence list (27), and Functional use (18). The search results are as follows:

Search Results
'plastic' found in the extracted text of 919 documents in 0.032 seconds.
PLASTIC MATERIAL Raw chemical name: plastic Number of products related to document: 1
NON-ASBESTOS GASKET MATERIAL,VARROIOUS SHAPE Raw chemical name: plastic Number of products related to document: 1
V6oRT,RECHARGEABLE BATTERY Raw chemical name: plastic Number of products related to document: 1
43027420 Raw chemical name: plastic Number of products related to document: 1
bptwt Raw chemical name: plastic Number of products related to document: 1
PRINT COATER APPLICATOR Raw chemical name: plastic Number of products related to document: 1
AC-207K CALCIUM LEAD BATTERY, WET Raw chemical name: plastic Number of products related to document: 1
MN 1400 (C) ALKALINE MANGANESE DIOXIDE CELL Raw chemical name: plastic Number of products related to document: 1

Figure 4.1: ChemExpo Search Box and Results Tabs

Table 4.1: Data fields searched for each Information Category within ChemExpo. The results of these searches are shown on the Search Result Tab for the appropriate Information Category

Information Category	Data Fields Searched
Documents	<ul style="list-style-type: none"> • Document title • Document subtitle • PMID • EPA regulation number • Raw chemical name • Raw CASRN • Curated chemical name • Curated CASRN
Products	<ul style="list-style-type: none"> • Product name • Brand name • Manufacturer • Product description
Product Use Categories	<ul style="list-style-type: none"> • PUC name • PUC description
Chemicals	<ul style="list-style-type: none"> • Raw chemical name • Raw CASRN • Curated chemical name • Curated CASRN • DTXSID
Chemical List Presence Keywords	<ul style="list-style-type: none"> • List Keyword name • List Keyword description
Function Categories	<ul style="list-style-type: none"> • FC name • FC description

Part I: Explore

When exploring data through ChemExpo there are few different types of pages that will be encountered. highlights the different ways to access the various pages that are described in more detail in the following section.

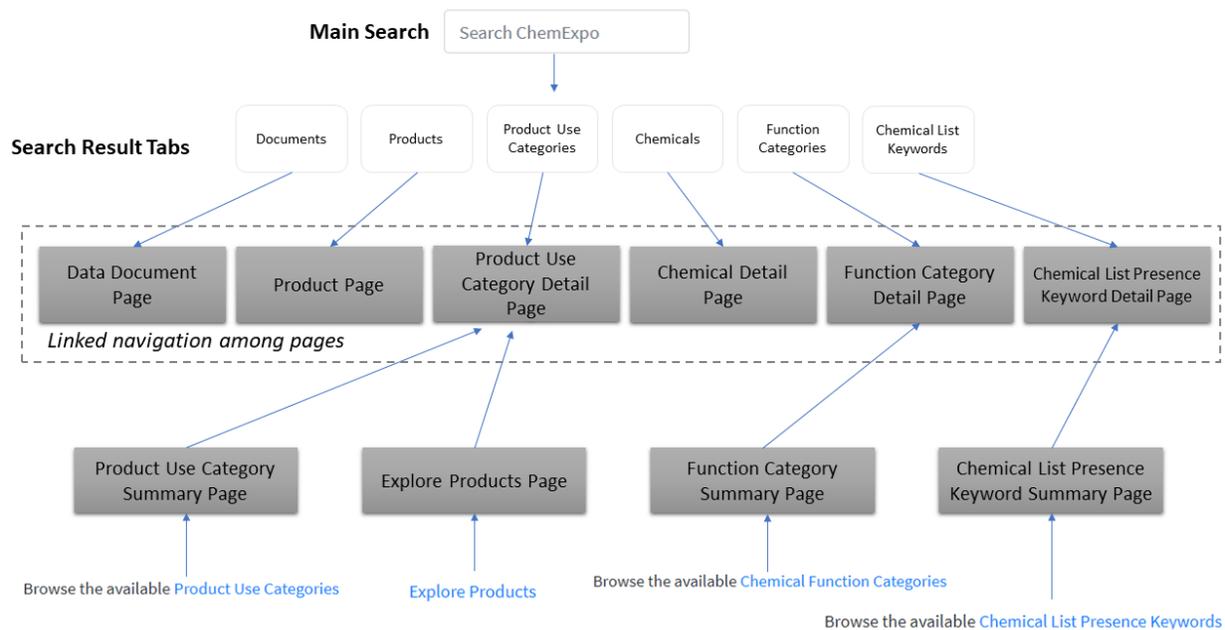


Figure I: Using the main search tab in ChemExpo allows users to explore all six information categories.

Chapter 5: Explore Chemicals

How do I identify what type of information is available for a specific chemical?

From the [ChemExpo homepage](#), identify a chemical of interest by typing in the chemical name, CASRN or DTXSID in the main search bar and pressing the Enter key. Navigate to the 'Chemicals' tab to view the resulting chemicals available in ChemExpo. Clicking on a specific chemical DTXSID takes the user to the Chemical Detail page.

Tip

If the DTXSID is known, Users can access the chemical page directly by replacing the DTXSID of interest within the page URL:

`https://ccte-chemexpo.epa.gov/chemical/DTXSID#####/.`

Users can view, filter, and download data related to a specific chemical of interest on the Chemical Detail page. The page displays any Product Use Categories (PUCs) whose products contain the searched chemical, as well as tables containing information on Products, Data Documents, Function Categories (FCs), and Chemical List Presence Keyword (List Keyword) Sets associated with a chemical. To view any data associated with a chemical of interest, use the Information Category Table Tabs at the bottom of the page. The details of each tab are described in [Table A.1](#).

5.1 Product Data

How do I filter the available information for a chemical to a specific Product Use Category (PUC?)

To filter the available information for a chemical to a specific PUC, use the PUC bubble plots ([Figure 5.1](#)). There are three bubble plots that organize PUCs by kind: Formulation, Article, and Industrial/Occupational PUCs. Click on the filter icon next to any PUC to filter the 'Products' 'Documents', and 'Functional Use' tables below the PUC bubble plots to show all entries related to the PUC of interest.

2-Butoxyethanol

DTXSID1024097

111-76-2

Explore products containing this chemical. Use the filter buttons to filter the tables below to certain Product Use Categories. Click the info button (i) next to each PUC to access its detailed chemical and product data.

Note: the high-level PUCs link to products that can't be further categorized; drill down to access products with refined PUC categorizations.

Formulation PUCs		
General Category - Product Family - Product Type		
> Arts and crafts/office supplies (i) ▾		22
> Cleaning products and household care (i) ▾		677
> Electronics/small appliances (i) ▾		17
> Food and drug (i) ▾		29
> Home maintenance (i) ▾		1113
> Landscape/yard (i) ▾		5
> Personal care (i) ▾		138
> Pesticides (i) ▾		9
> Sports equipment (i) ▾		6
> Vehicle (i) ▾		222

Article PUCs		
General Category - Product Family - Product Type		
Batteries (i) ▾		15
Cons. electronics, mech. appliances, and machinery (i) ▾		3
Construction and building materials (i) ▾		59
Furniture and furnishings (i) ▾		12
> Other direct contact consumer goods (i) ▾		2

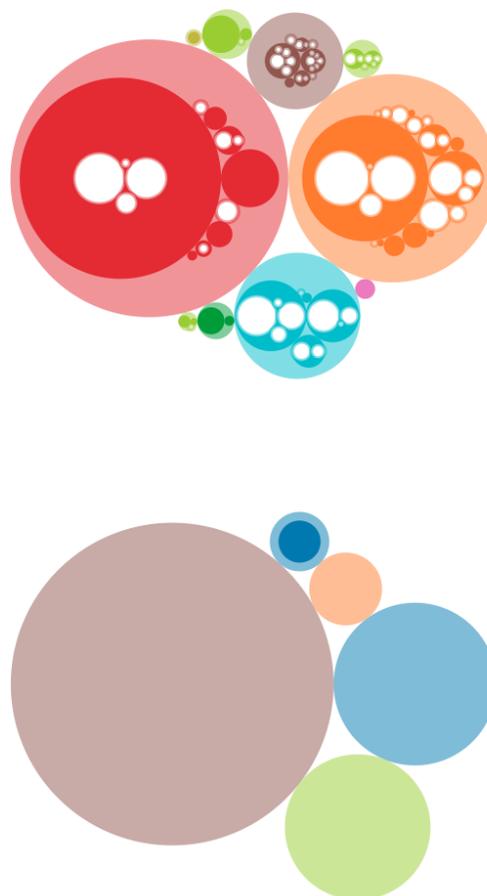


Figure 5.1: ChemExpo Chemical Detail Page; PUC Bubble Plot

(i) Note

When filtering on the coarsest PUC level (General Category) only products that have not been assigned a lower-level Product Family or Product Type will appear in the filtered table. Expand the PUC menu to the left of the bubble plots to identify products associated with more refined PUC levels.

How do I identify what specific consumer products have been associated with a chemical?

To identify what types of products have been associated with a chemical, view the 'Products' tab (Figure 5.2). The table displays all products containing the specified chemical of interest. For each product entry, the table lists the associated data document PUC, PUC Kind and PUC Classification method. The classification method describes how the product was assigned to the given PUC (e.g., via manual or automated assignment). Users can filter the table by PUC kind and Classification Method, for instance to find only 'Industrial/Occupational' use products. Users can also use the search field right above the table to filter the table based on data in any of the available columns. Users can also access (via clicking on product or document names) the individual Product or Data Document pages, as well as the detail page for any PUC in the table.

Associated List Presence Keyword Sets ▼

Products
Documents
Functional Uses

Products Containing "2-Butoxyethanol" Show All Products

Show entries Search:

Product	Document	PUC	PUC Kind	Classification Method
#410 graffiti remover	410 GRAFFITI REMOVER	Cleaning and safety - cleaning products - graffiti remover	Occupation	Manual
000088 thinner slow	000088 THINNER SLOW	Home maintenance - paint/stain and related products	Formulation	Automatic
011a_ aerosol cwf semi-gloss_ (see supplemental)	011A, AEROSOL CWF SEMI-GLOSS, (SEE SUPPLEMENTAL)			
031 crazy clean cleaner 19 oz	031 CRAZY CLEAN CLEANER 19 OZ	Cleaning products and household care - general household cleaning - surface cleaner	Formulation	Automatic
09-d0072-000 masterpiece penetrating wax stripper	09-D0072-000 MASTERPIECE PENETRATING WAX STRIPPER	Home maintenance - paint/stain and related products - stripper	Formulation	Manual Batch
09779awa gray aqua-zen enamel dark	09779AWA GRAY AQUA-ZEN ENAMEL DARK	Home maintenance - paint/stain and related products	Formulation	Automatic
1014 scotchgard brand cleaner for upholstery	1014 SCOTCHGARD BRAND CLEANER FOR UPHOLSTERY	Cleaning products and household care - upholstery specific - upholstery cleaner	Formulation	Manual Batch
10516 blair no odor spray fix	10516 BLAIR NO ODOR SPRAY FIX	Personal care - hair styling and care - hair spray	Formulation	Automatic
1249z_ heavy duty institutional formula (supp	1249Z, HEAVY DUTY INSTITUTIONAL FORMULA (SUPP DATA)			

Figure 5.2: ChemExpo Information Category Tables on Chemical Detail Page

5.2 Additional Chemical Information

How can I find additional information (e.g., chemical property or toxicity information) for a chemical reported in a specific consumer product?

To find additional chemical property or toxicity information related to a chemical of interest, access the CompTox Chemicals Dashboard. On the Chemical Detail page, the linked DTXSID below the chemical name header will take the user to an external EPA Website for

the public CompTox Dashboard (Figure 5.3). This chemical page provides links to additional information related to the chemical.

The CompTox Chemicals Dashboard is also linked from ChemExpo via the DTXSID on each chemical card on the Data Document page. For curated chemicals, users can click on the DTXSID to be redirected to the Dashboard chemical landing page.

The screenshot shows the CompTox Chemicals Dashboard for 2-Butoxyethanol. The page header includes navigation links (Home, Search, Lists, About, Tools), a search bar, and a 'Submit Comments' button. The main content area features a chemical structure of 2-Butoxyethanol, its name, CAS number (111-76-2), and DTXSID (1024097). A sidebar on the left provides a menu of categories: Chemical Details, Executive Summary, Physchem Prop, Env. Fate/Transport, Hazard, Safety > GHS Data, ADME > IVIVE, Exposure, Bioactivity, GenRA, Literature, Links, and Comments. The main content area is divided into several sections: Wikipedia (describing 2-butoxyethanol as an organic compound with formula $\text{BuOC}_2\text{H}_4\text{OH}$), Quality Control Notes, Intrinsic Properties (Molecular Formula: $\text{C}_8\text{H}_{18}\text{O}_2$, Average Mass: 118.176 g/mol, Monoisotopic Mass: 118.09938 g/mol), Structural Identifiers, Linked Substances, Presence in Lists, and Record Information. A URL bar at the bottom shows <https://comptox.epa.gov/dashboard/>.

Figure 5.3: CompTox Chemicals Dashboard

5.3 Chemical List Presence Data

How do I filter the documents associated with a chemical to those associated with a specific Chemical List Presence Keyword set?

To filter the documents associated with a chemical to those associated with a specific keyword set, use the List Keyword Sets filter. Click on the 'Associated List Presence Keyword Sets' banner (see Figure 5.2) below the bubble plots to expand the list of all keyword sets related to this chemical. Click on the filter icon next to any keyword set to filter the 'Documents' table below to show all Chemical Presence documents associated with that keyword set.

5.4 Function Data

How do I identify what types of Function Categories (FCs) have been associated with a chemical?

ChemExpo keeps track of two different, but related, functional uses for chemicals: their reported function and their harmonized Function Category (FC). Reported function is the exact text that the data source used to describe the function of a chemical substance. Harmonized Function Category is a curation of the reported function to the standardized list of FCs. To identify what reported functions have been associated with a chemical, use the 'Function Categories' tab (see Figure 5.2). The "Function" table shows all reported

functions and FCs associated with the chemical. Users can use the search bar above the table to filter it to a unique FC.

Chapter 6: Explore Products or Product Use Categories

6.1 Product Information in ChemExpo

What types of products are included in ChemExpo? Where can I find definitions for the different Product Use Categories (PUCs)?

Products in ChemExpo are categorized into Product Use categories (PUCs), which indicate the type of the product based on information provided in the original data source. More information can be found in Data Overview ([Chapter 3](#)). PUCs include consumer formulations, consumer articles, and industrial/occupational products. Definitions for the different product categories are found on the [PUCs Summary Page](#).

6.2 Browsing PUCs and their Data

How do I browse the available Product Use Categories (PUCs) and their data? Users can browse the available PUCs via the Explore Products tab (on the main ChemExpo menu bar) or via the PUCs Summary page.

1. Users can click on the 'Explore Products' tab on the main ChemExpo navigation bar, to view bubble plots for all available PUCs. These visualizations are organized by PUC Kind, with three separate plots for Formulation, Article, and Industrial & Occupational PUCs. Users can click on any General Category in the tables to the left of the bubble plot to expand the PUC list to display lower-level Product Families and Product Types within the General Category. This also zooms in the corresponding bubble plot visualization. Users can also click the info button next to each PUC to access its detailed chemical and product data.
2. Users can access the PUC Summary table through the 'Product Use Categories' link on the homepage or clicking on 'Get Data' on the top ChemExpo menu and then clicking 'Product Use Categories', or directly through <https://ccte-chemexpo.epa.gov/pucs/>. The PUCs summary table can be searched and filtered using the search bar above the table. The table can also be sorted and filtered on each column, by clicking on any table column header and entering in a filter query. Users can also click the blue 'View Detail' button to access the PUC Detail page for a PUC of interest.

6.3 PUC Detail Pages

[« Back to PUC list](#)

Cleaning products and household care

Kind Formulation

Description cleaning and household care products that can not be placed in a more refined category

Assumed Attributes

No assumed attributes are associated with this PUC

Allowed Attributes

aerosol exterior foam spray gel impregnated sheets
interior liquid paste powder pump spray solid
wipes

Additional Statistics

Products 1595
Cumulative Products 13843
Documents 1558
Curated Chemicals 707

Products Documents Chemicals Functions

Products containing chemicals assigned to "Cleaning products and household care"

Show 10 entries

Search:

[Download Products and Chemical Weight Fractions](#)

Product	Brand Name	Manufacturer	Classification Method
52666 general purpose cleaner		SAFETY-KLEEN CORP	Manual
51666 general purpose cleaner		SAFETY-KLEEN CORP	Manual
50666 general purpose cleaner		SAFETY-KLEEN CORP	Manual
11666 general purpose cleaner		SAFETY-KLEEN CORP	Manual
10667 general purpose cleaner		SAFETY-KLEEN CORP	Manual
10666 general purpose cleaner		SAFETY-KLEEN CORP	Manual
666 general purpose cleaner		SAFETY-KLEEN CORP	Manual
0145-6_vortex		ECOLINK INC	Manual
vortex		ECOLINK INC	Manual
classic		BRULIN/INDIANAPOLIS, IN 46206	Manual

Figure 6.1: ChemExpo PUC Details Page

What information is provided on the PUC Detail Page?

The PUC Detail page (Figure 6.1) displays additional information related to a specific PUC, including the Kind, definition, allowed attributes, assumed attributes, and additional statistics. Users can also view available data based on Products, Documents, Chemicals and Function Categories related to a Product Category of interest (see sections below). The details of each tab are described in Table A.2. The Products and Chemicals tabs for PUCs are described in more detail in (Section 6.4 and Section 6.5)

6.4 Finding Products via PUCs

How do I find information about specific products associated with a Product Use Category (PUC)?

Users can find information about specific products associated with a PUC of interest via the PUC Detail page. The PUC Detail page is accessed via the Explore PUCs or PUC Summary page, as described in [Section 6.3](#).

From the Products table, users can click on individual product entries to access more information about specific products on the Product page. There is also a Documents table on this page used to display the individual Data Documents associated with a PUC. Usually, there is a one-to-one relationship between a Document and Product but in some cases multiple products are defined in a single document. In this case, individual products are defined based on unique Universal Product Codes (UPCs) provided in the document. Specific products contain additional product details, such as brand name, manufacturer, and product description. Users can also access the linked document to access document metadata and chemical information. The PUC assigned to the product is also displayed on the Product page.

There are 5 different methods used to assign products to PUCs in ChemExpo. These are referred to as “PUC Classification Methods”. See [Table 6.1](#) for more detail on each Classification Method, and how each is used to assign products to PUCs. The product Classification Method appears alongside product entries on Product tables across ChemExpo. Classification Methods are ranked in terms of level of confidence in the assignment, based on manual or automated assignment, and how much information is available to the curator. Note that the application only displays the highest confidence level PUC assigned to a particular product.

Table 6.1: PUC Classification Methods, their descriptions, and the level of confidence assigned to each method. Confidence Levels range from 1 to 5 with 1 being the greatest Confidence Level and 5 being the lowest.

Confidence Level	PUC Classification Method	Code	Definition
1	Manual Assignment	MA	Manual assignment of an individual product to a PUC. Available on the Product page. This assignment method has the highest level of confidence because curators view the individual product and source document, and assign a category based on their understanding of the product details.
2	Automatic Conversion	AC	Conversion of AU assignments. In this process, curators review Automatic PUC Assignments and either ensure the assignment was correct (and assign AC) or unassign the AU PUC.

Confidence Level	PUC Classification Method	Code	Definition
3	Bulk Assign Product to PUC	MB	Assignment of products to PUCs based on Product title, Manufacturer, and Brand name keywords. This assignment method has medium confidence because curators assign multiple products at once based on the name, manufacturer, and brand.
4	Bulk Assign Raw Category to PUC	BA	Assignment of products to PUCs based only on the “recommended product use” data extracted from documents within a data group. Curators bulk assign products to a single PUC based on the Raw Category, without viewing individual products. This assignment method has a low level of confidence because curators are not viewing individual products and relies on general product use data provided by manufacturers.
5	Automatic	AU	Assignment of products to PUC using a machine learning model that predicts a PUC assignment for a product based on its name, brand, and manufacturer.

6.5 Chemical Data

How can I find all the chemicals that have been associated with a Product Use Category?

Users can find information about specific chemicals associated with a PUC of interest via the PUC Detail page. The PUC Detail page is accessed via the Explore PUCs or PUC Summary page, as described in the section above.

From the Chemicals table (accessed via the “Chemicals” tab in the center of the page), users can view all chemicals that have been curated to a specific PUC. The DTXSID column allows users to access the Chemical Detail page for more information on the specific chemical. The count is split into two columns, for Manual and Automatic counts. The Count indicates how many times the chemical appears in data documents that have been curated to the product use category, either via the manual classification methods, or via the automatic PUC classification method. The manual classification data are more reliable, but the automatic classifications allow more individual product documents to be identified as potentially relevant to a particular PUC. Users can refer to the product table on the PUC Detail page for data on the specific products that have been mapped in an automated manner.

Chapter 7: Explore Chemical Function

7.1 Browsing Chemical Functions

How do I browse the available Function Categories (FCs) and their data?

ChemExpo includes functional use data, as described in [Chapter 3](#), organized by harmonized Function Categories (FCs). Users can browse the available FCs via the Function Categories Summary page, available via the 'Browse available Function Categories' [link](#) on the homepage. The FC table on the Summary page can be searched and filtered using the search bar above the table. The table can also be sorted and filtered on each column, by clicking on any table column and entering in a filter query. The table displays a count of the number of products and number of documents associated with a given functional category. Users can also click the blue 'View Detail' button to access the Function Category Detail page for a function of interest.

[« Back to FC list](#)

pH regulating agent

Description Chemical substance used to alter, stabilize, or control the pH (hydrogen ion concentration) within a desired range. Also referred to as a buffering agent; pH adjuster; pH regulating agent; or neutralizing agent.

Additional Statistics

Products	2426
Documents	2914
Chemicals	295

Products	Documents	Chemicals
-----------------	------------------	------------------

Products containing chemicals assigned to "pH regulating agent"

Show 10 entries

Search:

[Download Products Associated with Function](#)

Product Name	Data Document	Product Use Categories	Classification Method
Laundry Detergent (powder)	laundry detergent powder 50 oz en	Cleaning products and household care - laundry and fabric treatment - laundry detergent	Manual
ARM & HAMMER w/OxiClean Powder Laundry Detergent- Fresh Scent	ARM & HAMMER w/OxiClean Powder Laundry Detergent- Fresh Scent	Cleaning products and household care - laundry and fabric treatment - laundry detergent	Manual
ARM & HAMMER Powder Laundry Detergent-Clean Burst	ARM & HAMMER Powder Laundry Detergents	Cleaning products and household care - laundry and fabric treatment - laundry detergent	Manual
ARM & HAMMER Powder Laundry Detergent-Clean Burst w/Color Safe Bleach	ARM & HAMMER Powder Laundry Detergents	Cleaning products and household care - laundry and fabric treatment - laundry detergent	Manual
ARM & HAMMER Powder Laundry Detergent-Perfume & Dye Free	ARM & HAMMER Powder Laundry Detergents	Cleaning products and household care - laundry and fabric treatment - laundry detergent	Manual
ARM & HAMMER Powder Laundry Detergent-Clean Mountain	ARM & HAMMER Powder Laundry Detergents	Cleaning products and household care - laundry and fabric treatment - laundry detergent	Manual
ARM & HAMMER Powder Laundry Detergent-Cool Breeze	ARM & HAMMER Powder Laundry Detergents	Cleaning products and household care - laundry and fabric treatment - laundry detergent	Manual
ARM & HAMMER Powder Laundry Detergent-Alpine Clean	ARM & HAMMER Powder Laundry Detergents	Cleaning products and household care - laundry and fabric treatment - laundry detergent	Manual
ARM & HAMMER ESSENTIALS Powder Laundry Detergent- Emerald Mist	ARM & HAMMER ESSENTIALS Powder Laundry Detergent	Cleaning products and household care - laundry and fabric treatment - laundry detergent	Manual
Toilet Bowl Cleaner	toilet bowl cleaner en	Cleaning products and household care - bathroom - bathroom cleaner	Manual

Figure 7.1: ChemExpo FC Detail Page

The Function Category Detail page (Figure 7.1) displays additional information related to a specific FC, including the description and additional statistics on available data. On the detail page users can also view available data for Products, Documents, and Chemicals related to the FC. These data are accessed via the three labeled tabs in the center of the page, each of which accesses a corresponding data table. Both composition documents and functional use documents can be associated with an FC. The Data Document pages, which hold the raw and curated chemical and function data, can be accessed in the Document table. The details of the tabs and data tables are described in Table A.3. The Products and Chemicals tabs for FCs are described in more detail in Section 7.2 and Section 7.3.

7.2 Chemical Data for a Functional Category

How do I find information about specific chemicals associated with a given Function Category (FC)?

Users can find information about specific chemicals associated a Function Category (FC) via the FC Detail page. The FC Detail page is accessed via the FC Summary page, as described in the section above (or via search results on the Function Category tab).

The Chemicals table on the FC Detail page (accessed via the “Chemicals” Tab in the center of the page) contains a list of chemicals associated with the FC. The DTXSID column allows users to access the Chemical Detail page for more information on the specific chemical.

7.3 Product Data for a Functional Category

How do I find information about specific products associated with a given Function Category (FC)?

Users can find information about specific products that contain a chemical with a specific Function Category (FC) via the FC Detail page. The FC Detail page is accessed via the FC Summary page, as described in the section above (or via search results on the Function Category tab).

The Products table on the FC Detail page (accessed via the “Products” Tab in the center of the page) contains a list of products containing a chemical associated with the FC. Users can click on individual product entries to access more information about specific products on the Product page or their associated documents or PUCs. Usually, there is a one-to-one relationship between a document and product but in some cases multiple products are defined in a single data document. Note that the Classification Method column indicates the manual or automatic method by which the product has been assigned to the listed PUC.

7.4 Interpreting FCs

How should I interpret the assignment of manufacturer-reported functional uses to Function Categories (FCs) in OECD?

Chemicals with manufacturer-reported functional uses are manually curated to FCs. Curators assign the text of a reported function to a harmonized category, so that every instance of a reported function is mapped to the same harmonized category. i.e., every instance of ‘viscosity controlling agent’ is mapped to the OECD function ‘viscosity modifier’. This removes any spelling or punctuation errors in the original reported text and allows for consistency in terminology across the application.

Chapter 8: Explore General Chemical Uses

8.1 Browse Chemical List Presence Keywords

How do I browse the available data on general chemical use via Chemical List Presence Keywords?

[← Back to LP list](#)

Drinking_Water

Media

water intended for drinking, or related to drinking water; includes bottled water, finished water from drinking water treatment plants, and untreated water that has been denoted as a drinking source

Filter By List Presence Keyword Sets Associated with Chemicals ▼

Documents Chemicals

Show entries Search: Download Documents Associated with Keyword

Document Title ↑↓	All Keywords Present in Document
2001 Pesticide Residues in Drinking Water	agricultural ; detected ; drinking_water ; pesticides ; residue
2002 Pesticide Residues in Drinking Water	agricultural ; detected ; drinking_water ; pesticides ; residue
2003 Pesticide Residues in Drinking Water	agricultural ; detected ; drinking_water ; pesticides ; residue
2004 Pesticide Residues in Water, Finished	agricultural ; detected ; drinking_water ; pesticides ; residue
2004 Pesticide Residues in Water, Untreated	agricultural ; detected ; drinking_water ; pesticides ; residue
2005 Pesticide Residues in Bottled Water	agricultural ; detected ; drinking_water ; pesticides ; residue
2005 Pesticide Residues in Water, Finished	agricultural ; detected ; drinking_water ; pesticides ; residue
2005 Pesticide Residues in Water, Untreated	agricultural ; detected ; drinking_water ; pesticides ; residue
2006 Pesticide Residues in Bottled Water	agricultural ; detected ; drinking_water ; pesticides ; residue
2006 Pesticide Residues in Water, Finished	agricultural ; detected ; drinking_water ; pesticides ; residue

Showing 1 to 10 of 62 entries Previous 1 2 3 4 5 6 7 Next

Figure 8.1: ChemExpo List Keyword Presence Detail Page

ChemExpo includes general chemical use data curated from public documents, as described in [Chapter 3](#), organized by Chemical List Presence Keywords (List Keywords). Users can browse the available List Keywords the Chemical List Presence Keywords Summary page ([Figure 8.1](#)), accessed via the 'Browse available Chemical Presence Keywords' [link](#) on the homepage. The Keyword table on the Summary page can be searched and filtered using the search bar directly above the table. The table can also be sorted and filtered on each column, by clicking on any table column header and entering in a filter query. The table displays List Keywords, their kinds and definitions, and a count of the documents associated with each. More information about the kinds of List Keywords is found in

Table 8.1. Users can also click the blue ‘View Detail’ button to access the detail page for any List Keyword.

Table 8.1: Description of Kinds of List Keyword Presence.

Kind	Definition
General use	Related to general chemical use
PUC – article	Keyword is a product use category (PUC) of kind “article”
PUC – formulation	Keyword is a product use category (PUC) of kind “formulation”
PUC – industrial	Keyword is product use category (PUC) of kind industrial/occupational
Location	Related to location/origin of the document/list
Manufacturing	Related to the manufacturing process
Foods & Agriculture	Related to food and agriculture
Specialty list	Keyword refers to a recognized, specialty list of chemicals, e.g., a regulatory list
Subpopulation	Keyword denotes a specific population associated or affected by the document/list
Media	Associated with a specific environmental media, e.g., as in a measurement study
Modifiers	Keyword modifies other keywords assigned to the same list/chemical combination

8.2 Data Documents

How do I find information about specific documents associated with a given Chemical List Presence Keyword set?

Users can find information about Data Documents associated with List Keyword and List Keyword sets (specific Keyword combinations) of interest via the Chemical List Presence Keyword Detail page. The detail page is accessed via the List Presence Keyword Summary page, as described in [Section 8.1](#).

The Chemical List Presence Keyword Detail page displays additional information related to a specific keyword, including a table containing all the available Data Documents that contain a given Keyword, and a table of all chemicals associated with the Keyword. There is a filter functionality above the tables that allows a user to filter the table to the selected Keyword Sets. The table details are included in [Table A.4](#).

Chapter 9: Explore Data Documents

9.1 Accessing Data Documents

glass + surface cleaner waterfall

method products, phc

Reported Chemical Name	Reported CAS	Curated Chemical Name	Curated CAS	DTXSID	Count
Reported Chemical Name: water (aq)	Reported CAS: 7732-18-5	Curated Chemical Name: Water	Curated CAS: 7732-18-5	DTXSID6026296	1
Reported Chemical Name: ethanol	Reported CAS: 64-17-5	Curated Chemical Name: Ethanol	Curated CAS: 64-17-5	DTXSID0020584	2
Reported Chemical Name: sodium carbonate	Reported CAS: 497-19-8	Curated Chemical Name: Carbonic acid sodium salt (1:2)	Curated CAS: 497-19-8	DTXSID1029621	3
Reported Chemical Name: fragrance (parfum)	Reported CAS: n/a	Curated Chemical Name:			4
Reported Chemical Name: decyl glycoside	Reported CAS: 68515-73-1	Curated Chemical Name: D-Glucopyranose, oligomeric, decyl acryl glycos...	Curated CAS: 68515-73-1	DTXSID70872594	

ChemExpo Data Document Page for a Product

How do I access specific Data Documents?

There are many ways to access documents in ChemExpo. Users can access Data Documents directly from the 'Documents' tab on the Chemical, Product Use Category, Function Category, and Chemical List Presence Keyword Detail pages (described in sections [Section 5.3](#), [Section 6.3](#), [Section 6.4](#), [Section 7.1](#), [Section 8.1](#)).

Users can also find documents via the main search. From the ChemExpo homepage, identify a document of interest by typing in any search keyword or text string in the main search bar and pressing the Enter key. Navigate to the 'Documents' Search Result Tab to view the resulting documents matching the searched term. Clicking on a specific document title takes the user to the Data Document page. [Section 3.5](#) provides an overview of the Data Document page for each type of document. For details of the data on the Data Document page, see [Table A.5](#).

9.2 Original Data

How do I access the original data source for a document?

ChemExpo does not provide direct access to the original source documents due to copyright laws. There are two tools in ChemExpo to assist users in accessing the original data source for a document, via an external URL and an external search tool.

- **View Source Document (External):** On the Data Document page, users can click the 'View Source Document (External)' button in the top-right corner to access the page URL obtained from the original reporting organization or manufacturer. These

URLs are captured during data extraction and may degrade over time. This button may be missing if the original source URL is unavailable.

- **Search for Source Document (External)**: On the Data Document page, users can click the 'Search for Source Document (External)' button in the top-right corner to be redirected to a Google search query, which includes the data document title and organization. These keywords are useful in finding the original document, if it is still accessible on the web.

Part II: Download

There are many locations to download both comprehensive data summaries and specific smaller datasets in ChemExpo. Generally, users can find specific data to download on Chemical, Product, Functional Category, Product Use Category, and List Presence Keyword Detail Pages. These downloads are found above the Information Category tables at the bottom of each of these pages.

For users who prefer data that cover the entirety of an Information Category within ChemExpo, the "Get Data" page offers the ability to download comma-separated value (csv) files with the following data:

- All currently released data within an Information Category
- Metadata, where available, related to the categories within an Information Category.

ChemExpo Get Data Explore Products Search factotum ? i User Guide Contact

Get Data

Download Metadata	
Download PUCs	Product Use Categories
Download PUC Attributes	
Download List Presence Keywords	List Presence Keywords
Download Function Categories	Function Categories

Download Bulk Data
Bulk Download List Presence Data
Bulk Download Function Data
Bulk Download Composition Data

Figure II: ChemExpo's Get Data page.

Specific Data

9.1 Chemical Data

How do I download data associated with a particular chemical?

To download data associated with a chemical, use the download buttons located in the top-right corner of the tables on the Chemical Detail page ([Chapter 5](#)).

- **Download Products and Weight Fractions:** returns a csv file with all products, product use categories, and chemical composition/weight fraction data available for a given chemical.
- **Download Function Categories:** returns a csv file with all data documents that contain functional use information available for a given chemical, including reported functions and harmonized function categories.

9.2 Product Use Category Data

How do I download data associated with a particular Product Use Category (PUC)?

To download data associated with a PUC, use the download buttons located in the top-right corner of the tables on the Product Use Category Detail page ([Section 6.3](#)). Users can download data for all Products, Chemicals, or Function Categories associated with a Product Use Category.

- **Download Products and Chemical Weight Fractions:** returns a csv file with all products, product use data, and chemical composition/weight fraction data available for a given Product Use Category.
- **Download Chemicals:** returns a csv file with all unique chemicals, CASRNs and DTXSIDs associated with a given Product Use Category.
- **Download Function Categories:** returns a csv file with all data documents that contain function category information categorized to a Product Use Category, including the specific chemical, and its reported function and harmonized Function Categories.

9.3 Function Category Data

How do I download data associated with a particular Function Category (FC)?

To download data associated with a FC, use the download buttons located in the top-right corner of the tables on the Function Category Detail page ([Section 7.1](#)). Users can download data for all Products or Chemicals associated with an FC.

- **Download Products:** returns a csv file with all products and Product Use Category data available for a given Function Category.
- **Download Chemicals:** returns a csv file with all unique chemicals, CASRNs and DTXSIDs associated with a given Function Category.

9.4 Chemical List Presence Data

How do I download data associated with a particular Chemical List Presence Keyword?

To download data associated with a Chemical List Presence Keyword, use the download buttons located in the top-right corner of the tables on the Chemical List Presence Keyword Detail page. Users can download data for all Documents or Chemicals associated with a keyword.

- **Download Documents:** returns a csv file with all Document and Keyword Set data available for a given keyword.
- **Download Chemicals:** returns a csv file with all unique chemicals, CASRNs, and DTXSIDs associated with a given Keyword.

9.5 Document Data

How do I download data associated with a particular document?

To download data associated with a specific Data Document, use the 'download chemical data' button located in the top-right corner of the Data Document page ([Section 9.1](#)). The csv download returned contains document metadata, as well as chemicals and any data associated with the chemical (composition, Function Category, Chemical List Presence Keywords).

Chapter 10: Bulk Data

10.1 Product Data

10.1.1 PUC Metadata

How do I download names and definitions for all PUCs available in ChemExpo?

From the “Get Data”, click on the Download PUCs button in the Download Metadata table.

10.1.2 PUC Attributes

How do I download the names and definitions data for PUCs?

From the “Get Data”, click on the Download PUC Attributes button in the Download Metadata table.

10.1.3 Product Composition

How do I download information about all products and their ingredients in ChemExpo?

From the “Get Data”, click on the Bulk Download Composition Data button in the Download Bulk Data table.

10.2 Function Category Data

10.2.1 FC Metadata

How do I download information on all FCs available in ChemExpo? From the “Get Data”, click on the Download Function Categories button in the Download Metadata table.

10.2.2 FC Data

How do I download information about all chemicals with functional use information in ChemExpo?

From the “Get Data”, click on the Bulk Download Composition Data button in the Download Bulk Data table.

10.3 General Use Data

10.3.1 List Presence Keywords Metadata

How do I download information on all Chemical List Presence Keywords available in ChemExpo?

From the “Get Data”, click on the Download List Presence Keywords button in the Download Metadata table.

10.3.2 List Presence Keywords Data

How do I download information about all chemicals associated with List Presence Keywords in ChemExpo?

From the “[Get Data](#)”, click on the Bulk Download List Presence Data button in the Download Bulk Data table.

Chapter 11: Glossary

Term	Information Category	Definition
Allowed Attribute	Product	An attribute that may be applied to a product categorized into a certain PUC. Allowed attributes are a subset of attribute choices associated with a certain PUC, which are manually assigned to a given product.
Article	PUC Kind	A manufactured item which is formed to a specific shape or design during manufacture that determines its end-use function.
Assumed Attribute	Product	An attribute that is automatically applied to any product categorized into a certain PUC. For example, all products placed in the 'children's toys' PUC will automatically be assigned the attribute 'child'.
Attribute	Product	Attributes of individual products that may be useful for estimating exposure. Attributes may describe the form of the product (e.g., spray) or population exposed (e.g., child). For occupational PUCs, attributes can refer to the specific industry.
Author	Document Metadata	The name of the author of an article, report, etc., as obtained from the original source.
Automatic Conversion	PUC Classification Method	Conversion of model predicted PUCs to a manual prediction. Curators verify that products assigned via the AU method belong in a given PUC and assign them to the PUC through this higher classification method.
Automatic Assignment	PUC Classification Method	Automated (natural language processing-based) assignment of products to PUC using a machine learning model that predicts a PUC assignment for a product based on its name, brand, and manufacturer.
Brand	Product	Typically, a name that can distinguish between one manufacturer's or seller's product from that of another, similar product sold or made by another manufacturer or seller.

Term	Information Category	Definition
Bulk Assign Product to PUC	PUC Classification Method	Assignment of products to PUCs based on Product title, Manufacturer, and Brand name keywords. This assignment method has medium confidence, as curators assign multiple products at once based on the name, manufacturer, and brand.
Bulk Assign Raw Category to PUC	PUC Classification Method	Assignment of products to PUCs based only on the "Raw Category" data extracted from documents within a data group. The 'Connect PUC' functionality allows curators to Bulk assign products to a single PUC based on the Raw Category, without viewing individual products. This assignment method has a lower level of confidence because curators are not viewing individual products and relies on the accuracy of the general product use data provided by manufacturers.
Chemical List Presence (CP)	Data Group Type	Chemical list presence data groups include general reports and similar documents providing general information on the use of chemicals. Documents may come from international, federal, or state agencies, trade associations, or other reputable sources. Chemical list presence documents are wide ranging in format, authorship group, and information provided. Quantitative information is not typically extracted from these documents, rather lists of chemicals identified in the document may be tagged with keywords allowing users to both identify general uses of chemicals, and to identify documents of interest more efficiently for a particular research effort. Functional use of chemicals may be identified in the document.
Classification Method	Product	Curation method, whether manual or automated (natural language processing-based), used to assign product to PUC (Product Use Category) [Manual, Automatic Conversion, Manual Batch, Bulk Assign, Automated].

Term	Information Category	Definition
Cleaned Weight Fraction	Composition	Cleaned weight fraction data. Raw Composition data is curated to a harmonized form (decimal fraction between 0-1).
Component	Composition or Chemical Presence	Used to separate reported chemicals based on components, as reported in the source document. For example, by product component, for products with multiple parts or items in a set.
Composition (CO)	Data Group Type	The Composition data group type includes documents providing information on the composition of products, typically produced by the product manufacturers. Typically includes information on chemicals contained within the product, and may include information on the weight fraction of these chemical ingredients, and the functional use of each ingredient. These documents are in the form of a Safety Data Sheet (SDS), Material Safety Data Sheet (MSDS), or similar. Composition documents are typically linked to Factotum products, which may be classified into a product use category (PUC).
Curated CAS	Chemical	Curated Chemical CAS Registry Number (CASRN) associated with DTXSID, provided by DSSTOX. The curated CAS may be different than the reported CAS.
Curated Chemical Name	Chemical	Curated Chemical Name associated with DTXSID, provided by DSSTOX. The curated chemical name may be different than the reported chemical name.
Data Document	Document Metadata	The individual documents representing the original sources of data stored within ChemExpo. If the data document is a journal article or report, this might be the title of the article or report; if the data document is an MSDS sheet, this may be the name of the product represented in the MSDS sheet. If no title is provided when the data document is registered by the curator, the title defaults to be the same as the file name of the data document.

Term	Information Category	Definition
Data Document Subtitle	Document Metadata	The document subtitle. May also contain information specific to the document table or figure extracted.
Data Document Title	Document Metadata	The title of the data document. e.g., if the data document is a journal article or report, this might be the title of the article or report; if the data document is an MSDS sheet, this may be the name of the product represented in the MSDS sheet. If no title is provided when the data document is registered by the curator, the title defaults to be the same as the file name of the data document.
Data Document Type	Document Metadata	Indicates the type of document format, as obtained from the original source. i.e., MSDS, journal article, governmental Report.
Data Group	Document Metadata	The second level of organization for documents. A data group belongs to only one data source, and a document may belong to only one data group. All documents within the data group belong to the same data group type.
Data Group Type	Document Metadata	An internal curation classifier used to group documents that contain the same type of information. E.g., composition, functional use, chemical presence.
Data Source	Document Metadata	The highest-level organizational category which groups data documents that may come from a single source, have the same data type, or have other characteristics in common. Data source name may refer to the organization supplying the data (e.g., Walmart) or the organization which collected the data (e.g., Danish EPA).

Term	Information Category	Definition
Detected Flag	Chemical Presence	Indicates whether a chemical was measured or identified in environmental media or consumer products. "Yes" means that the chemical was detected by the study, while "No" could mean that the chemical was tested for, but not found; below the limit of detection for the instrumentation used; or instances where a measurement was not reported due to issues with measuring equipment, QA concerns, or other reason.
Document Date	Document Metadata	The date the document was created or last updated, as described on the original source document.
Document URL	Document Metadata	External URL that links to the original document or page the document was downloaded from. URLs may degrade over time but were active at the time the data was originally accessed.
DOI	Document Metadata	Digital Object Identifier (DOI) used to permanently identify an article or document and provide an external link.
DTXRID	Chemical	DSSTox Record Identifier. Unique identifier for each chemical record, provided by DSSTOX.
DTXSID	Chemical	DSSTox Substance Identifier. Unique substance identifier, provided by DSSTOX.
EPA Registration Number	Document Metadata	EPA-assigned identifier, applies to documents about specific pesticides or disinfectants.
Formulation	Product Kind	Formulation based consumer products (e.g., liquids, gels, powders etc) A formulated product is composed of at least two ingredients which are selected, processed and combined in a specific way to obtain well-defined target properties, functionality and performance.

Term	Information Category	Definition
Functional Use (FU)	Data Group Type	Functional use data groups include documents where potential functional uses (i.e., the purpose the chemical serves/may serve) are the only information contained within in the document. Documents may be sourced from chemical manufacturers or other organizations. If the functional use of a chemical is provided in association with the chemical's use in a particular product, or the document contains composition data, it should be classified within a composition data group type. Similarly, if other information about the use of the chemical (in addition to functional use) is provided in the document, the document should be classified within a chemical list presence data group type.
Functional Use Category	Functional Use	Harmonized identifiers describing the function a specific chemical serves (for example: fragrance, cleaning agent, filler, etc.) in a product. Functional use categories are obtained from OECD or developed by EPA, and are assigned via manual curation.
HERO ID	Document Metadata	Reference ID for EPA's Health & Environmental Research Online (HERO) Database.
Industrial/Occupational	Product Kind	Specialty products generally intended for industrial or occupational use.
Ingredient Rank	Chemical Composition	The numerical order the chemical ingredient appears in an ingredient list.
List Presence Keyword	Chemical Presence	Descriptive keywords assigned to chemical records, that are used to convey information about general chemical use.
List Presence Keyword Kind	Chemical Presence	Organizational system that identifies broader categories that group related keywords. List Presence Keywords are grouped Kind, which are categories that describe the keywords within (e.g., Location, PUC, Media, etc.).

Term	Information Category	Definition
List Presence Keyword Set	Chemical Presence	Refers to the group of chemical presence keywords assigned to a given chemical that appears in a data document. Descriptive keyword sets are used to convey information about general chemical use, and should be interpreted as a whole.
Manual Assignment	PUC Classification Method	Manual assignment of an individual product to a PUC. Available on the Product page. This assignment method has the highest level of confidence, as curators view the individual product and source document, and assign a category based on their understanding of the product details.
Organization	Document Metadata	Organization associated with the document. Can be the journal, manufacturer, organization that published a report, governing body, etc.
PMID	Document Metadata	The PubMed ID, provided if the source document is an article in the PubMed library.
Product	Product	A data record associated with a Chemical Composition document that contains product-relevant information such as manufacturer, brand, UPC, size, item description, product images, etc. There may be multiple products associated with a single composition document, as products are based on unique Universal Product Codes (UPC).
Product Attribute	Product	Tags describing qualities of a particular product, such as the product form (i.e., aerosol, liquid, gel), intended user demographic (i.e., child), special features of the product (i.e., two-component), or micro-environment in which the product is used (i.e., Exterior, Interior).
Product Kind	Product	Terms that organize PUCs based on the regulatory guidelines regarding the manufacture and ingredient reporting requirements for the products.

Term	Information Category	Definition
Product Use Category	Product	Identifiers assigned to products in Factotum, indicating the type of product assigned to each data record based on information provided in the original data source.
Provisional Flag	Chemical	Indicates whether the chemical record was provisionally assigned a DTXSID chemical identifier [Yes/No].
PUC General Category	PUC Hierarchy	Highest level in PUC Hierarchy. This is the most General Category describing a product's use. (i.e., personal care).
PUC Hierarchy	Product	Three-tiered organizational hierarchy for PUCs that consists of three levels, which ranges from general to more specific product types: General Category, Product Family, and Product Type.
PUC Product Family	PUC Hierarchy	Second Level in PUC Hierarchy. The Product Family describes a more specific type of product within the general category. (i.e., dental care, body hygiene, make-up and related).
PUC Product Type	PUC Hierarchy	Third (most specific) Level in PUC Hierarchy. The Product Type contains a specific type of product within a Product Family (i.e., toothpaste, mouthwash).
Raw Composition	Composition	Raw weight fraction/composition data, as reported on the source document. May be reported as percent, decimal fraction, or other form.
Reported CAS	Chemical	Chemical CAS Registry Number (CASRN) as provided on original source document.
Reported Chemical Name	Chemical	Chemical name as provided on original source document.
Reported Functional Use	Functional Use	Functional use of the chemical in a product, as reported on the original source document.
Revision Number	Document Metadata	The Revision or Version number as reported on original source.
UberPUC	Product	System for displaying products in PUCs based on Confidence in the Classification method.

Term	Information Category	Definition
Unit Type	Composition	The unit type for the concentration, as reported in source document.
Universal Product Code (UPC)	Product	A twelve digit code that uniquely identifies a product; typically found on the product's barcode.
Weight Fraction Type	Composition	Reported (Default) refers to weight fractions calculated using composition data present on the source document, while Predicted values are generated via model [predicted or reported].

Chapter 12: Frequently Asked Questions (FAQs)

12.1 Data and Maintenance Issues

Why can't ChemExpo provide the original Data Documents?

Data documents can't be directly provided by ChemExpo without copyright infringement. However, ChemExpo provides on the data document page a link to the original URL from which the document was obtained. PubMedIDs and DOIs are also provided where they are available.

Why can't the data document be found at the link provided?

The location at which an organization or a manufacturer provides a document may change over time. ChemExpo also provides a link to a Google Search for the document title to aid users in locating the document if it is still available elsewhere.

How often is ChemExpo updated?

The ChemExpo application will have updates as needed and as resources become available. The data in ChemExpo, however, is expected to be updated bi-annually.

What kind of updates are we planning in the future?

Updates to the underlying data in ChemExpo are planned as more information is collected and curated. Based on beta testing, some updates to the application itself may be planned based on available resources.

Where are the exposure predictions related to these chemicals?

Exposure predictions for chemicals can be found on the [U.S. EPA's CompTox Chemicals Dashboard \(CCD\)](#).

Can I suggest a new source for Curation in ChemExpo?

Yes, recommendations can be sent to the ChemExpo Curation Team via the [contact email](#). Data sources are prioritized for curation according to uniqueness (compared to other information in ChemExpo) and relevance to EPA.

Does ChemExpo contain information on exposure routes of specific products or product categories?

While ChemExpo does not contain this information specifically, the Product Use Categories (PUCs) contain attributes (for example, “spray”) which may allow mapping of a product category to specific exposure routes. In addition, mapping of formulation PUCs to specific release scenarios were provided in the Supplement of Isaacs et al. (2020) (although the PUCs used here have evolved since that publication).

CPDat also contains predicted information on chemical function from Quantitative Structure-Use Relationships (QSURs). Are these predictions available via ChemExpo?

Since these are model predictions as opposed to reported use information from public documents, these data are not included in ChemExpo. They are provided on the Exposure Tab of the CompTox Chemicals Dashboard.

Why are there no chemicals cards for certain products?

Many chemical ingredients in ChemExpo are extracted from safety data sheets (SDS) reported by the manufacturer (see [Appendix C](#)). However, these documents are only required to include ingredients for products that might pose a hazard to humans. If there are no substances that meet those requirements, there will be no chemical cards for that product. Future versions of ChemExpo may be able to filter out these products.

Why are there duplicate product entries in ChemExpo?

In the majority of cases, seemingly duplicate products are likely multiple products (unique UPCs) that have the same name. These documents are currently considered unique in the ChemExpo data model. The same products could have been collected from different data sources. In addition, some products have their formulations updated over time, when this happens there may be duplicate records of the same product, but from different years.

12.2 Usability Questions

How can I search for Safety Data Sheets (SDSs) online from ChemExpo?

Persistence of a web site or an online document over time is a common issue in all data curation efforts, ChemExpo makes every effort to ensure that we provide the URL of the source document from the time it was downloaded for curation. However, that does not mean that the URL will work in perpetuity.

Further, ChemExpo does not provide the actual SDS published by a company since that is copyrighted material. It does however, provide a URL to the site from which the SDS was downloaded. From a Data Document Details Page, the URL for a data source (at the time it was downloaded) is provided in the upper right hand corner as an external link button (square with an arrow pointing out of it). Alternatively, there is the function to automatically perform a search via the Google search engine for that Data Document title. This can be done by clicking the magnifying glass button in the upper right corner of the Data Document Details Page.

What are Universal Product Codes (UPCs) and Stub Entries within ChemExpo?

A Universal Product Code (UPC) is a unique number assigned to a product. This is typically found on a product's barcode. For some products within ChemExpo, the UPC is known. For others, however, the UPC is not known. In these cases, a stub identifier is assigned to products as a UPC code.

What if my ChemExpo search returns no results?

If there are no results from a ChemExpo search in any of the six information categories, this means that there is no curated information available that matches these terms in ChemExpo's underlying database. It does not necessarily mean the data does not exist, just that it is not currently contained within ChemExpo. If you know of a source that contains information you are looking for and would like to see it added to ChemExpo, please email the ChemExpo team at chemexpo.support@epa.gov.

How can a consumer use ChemExpo?

At its heart ChemExpo was developed to aid in providing information for performing exposure evaluations of chemicals. However, consumers are welcome to use this application to identify various chemical ingredients noted to be in specific product types (product composition and use), why those chemicals may be there (functional use), and what general categories are associated with a chemical or document (list presence keywords). Consumers should be aware of the limitations of these data and their uses as described in the Disclaimer.

Why do some products contain chemicals stating “proprietary” ingredients?

Under the Hazard Communication Standard, manufacturers are allowed to claim ingredients and the composition in products as proprietary. Often when users find a chemical with the name “Proprietary” this means the ingredient’s name was not released on the SDS for a product. Other ingredients may be reported in general terms, for example, “fragrance”. ChemExpo reports ingredients as they were found on the original document.

References

Dionisio, Kathie L., Alicia M. Frame, Michael-Rock Goldsmith, John F. Wambaugh, Alan Liddell, Tommy Cathey, Doris Smith, et al. 2015. "Exploring Consumer Exposure Pathways and Patterns of Use for Chemicals in the Environment." *Toxicology Reports* 2: 228–37. <https://doi.org/10.1016/j.toxrep.2014.12.009>.

Dionisio, Kathie L., Katherine Phillips, Paul S. Price, Christopher M. Grulke, Antony Williams, Derya Biryol, Tao Hong, and Kristin K. Isaacs. 2018. "The Chemical and Products Database, a Resource for Exposure-Relevant Data on Chemicals in Consumer Products." *Scientific Data* 5 (1). <https://doi.org/10.1038/sdata.2018.125>.

Directorate, Environment. 2017. "Internationally Harmonised Functional, Product and Article Use Categories." Organisation for Economic Co-operation; Development.

Grulke, Christopher M., Antony J. Williams, Inthirany Thillanadarajah, and Ann M. Richard. 2019. "EPA's DSSTox Database: History of Development of a Curated Chemistry Resource Supporting Computational Toxicology Research." *Computational Toxicology* 12 (November): 100096. <https://doi.org/10.1016/j.comtox.2019.100096>.

Isaacs, Kristin K., Kathie Dionisio, Katherine Phillips, Charles Bevington, Peter Egeghy, and Paul S. Price. 2020. "Establishing a System of Consumer Product Use Categories to Support Rapid Modeling of Human Exposure." *Journal of Exposure Science and Environmental Epidemiology* 30 (1): 171–83. <https://doi.org/10.1038/s41370-019-0187-5>.

Koval, Lauren E., Kathie L. Dionisio, Katie Paul Friedman, Kristin K. Isaacs, and Julia E. Rager. 2022. "Environmental Mixtures and Breast Cancer: Identifying Co-Exposure Patterns Between Understudied Vs Breast Cancer-Associated Chemicals Using Chemical Inventory Informatics." *Journal of Exposure Science and Environmental Epidemiology* 32 (6): 794–807. <https://doi.org/10.1038/s41370-022-00451-8>.

Phillips, Katherine A., John F. Wambaugh, Christopher M. Grulke, Kathie L. Dionisio, and Kristin K. Isaacs. 2017. "High-Throughput Screening of Chemicals as Functional Substitutes Using Structure-Based Classification Models." *Green Chemistry* 19 (4): 1063–74. <https://doi.org/10.1039/c6gc02744j>.

U.S. Occupational Health and Safety Administration. 2012. "Hazard Communication." U.S. Department of Labor. <https://www.osha.gov/laws-regs/regulations/standardnumber/1910/1910.1200>.

Part III: Appendices

Appendix A: Visualization Page Details

A.1: Chemical Detail Page Tables

Table A.1: Description of Tables on the Chemical Detail Page.

Table Name	Description	Table Columns	Column Description
Products	The products table features a list of all products contain a specific chemical.	Product	Product title, click the link to view the Product page.
		Document	Data document name, click the link to view the Data Document page.
		PUC	Product Category assigned to the product, click the link to view the Product Category Detail page.
		PUC Kind	The kind of Product Use Category, based on the regulatory guidelines regarding the manufacture and ingredient reporting requirements for the products.
Documents	The Documents table shows all composition data documents that contain a specific chemical.	Classification Method	Shows the classification method of a product (how the product was assigned to the specified PUC).
		Document	Data Document name, click the link to view the data document page.
		Data Type	Data type for the document. The type corresponds to the document format and type of information captured from each document.
		Document Date	The date the document was created or last updated, as described on the original source document.

Table Name	Description	Table Columns	Column Description
Function Categories	The Function Category table shows all documents that contain functional use data related to a specific chemical.	Data Type	Data type for the document. The type corresponds to the document format and type of information captured from each document. Functional Use data appears in Composition, Functional Use and Chemical Presence List type documents.
		Chemical Name	Chemical name associated with original data source, or curated DTXSID record, as applicable.
		Document	Data document containing the specific functional use associated with this chemical, click the link to view the data document page.
		Reported Functional Use	Functional use of the chemical in a product, as reported by the original data source.
		Harmonized Functional Use	Functional Category assigned via manual curation.

A.2: PUC Detail Page Tables

Table A.2: Description of Tables on the PUC Detail Page.

Table Name	Description	Table Columns	Column Description
Products	The products table features a list of all products that have been assigned to a specific PUC.	Title	Product title, click the link to view the Product page.
		Brand Name	Product detail extracted for the brand name (if applicable).
		Classification Method	Shows the classification method of a product (how the product was assigned to PUC).
Documents	The Documents table shows all composition data documents related to a specific PUC.	Document	Data document name, click the link to view the data document page.

Table Name	Description	Table Columns	Column Description
Chemicals	The Chemicals table shows all chemicals found within data documents related to a specific PUC.	Data Type	Data type for the document. The type corresponds to the document format and type of information captured from each document. All documents will have the "Composition" data type.
		Document Date	The date the document was created or last updated, as described on the original source document.
		DTXSID	DSSTox substance identifier (DTXSID) used on the EPA CompTox Chemicals Dashboard, links to the Chemical Detail page.
		Curated CAS	Curated CAS Registry Number Associated with DTXSID.
		Curated Chemical Name	Curated Chemical Name Associated with DTXSID.
Function Category	The Function Category table shows all functional use data documents related to a specific PUC.	Count	Number of times chemical appears in data documents associated with this PUC.
		Document	Data document containing the specific chemical and functional use associated with this PUC, click the link to view the data document page.
		Chemical Name	Chemical name associated with original data source, or curated DTXSID record, as applicable.
		CAS	CAS Registry Number associated with original data source, or curated DTXSID record, as applicable.
		Reported Function Category	Functional use of the chemical in a product, as reported by the original data source.

Table Name	Description	Table Columns	Column Description
		Harmonized Function Category	Functional use category assigned via manual curation.

A.3: FC Detail Page Tables

Table A.3: Description of Tables on the Function Categories (FC) Detail Page.

Table Name	Description	Table Columns	Column Description
Products	The products table features a list of all products containing chemicals that have been assigned to a specific functional use category.	Product Name	Product title, click the link to view the Product page.
		Data Document	Associated Data document name, click the link to view the data document page.
		Product Use Categories	Product Category assigned to the product, click the link to view the Product Category Detail page.
Documents	The Documents table shows all data documents related to a specific functional use category.	Classification Method	Shows the classification method of the product (how the product was assigned to the specified PUC).
		Data Type	Data type for the document. The type corresponds to the document format and type of information captured from each document. FC data appears in Composition, FC and Chemical Presence List type documents.
		Document	Data document name, click the link to view the Data Document page.
		Document Date	The date the document was created or last updated, as described on the original source document.
		Reported FC	Functional use of the chemical in a product, as reported by the original data source.

Table Name	Description	Table Columns	Column Description
Chemicals	The Chemicals table shows all chemicals found within data documents related to a specific FC.	DTXSID	DSSTox substance identifier (DTXSID) used on the EPA CompTox Chemicals Dashboard, links to the Chemical Detail page.
		Curated CAS	Curated CAS Registry Number Associated with DTXSID, links to the Chemical Detail page.
		Curated Chemical Name	Curated Chemical Name Associated with DTXSID.

A.4: List Presence Keyword Detail Page Tables

Table A.4: Description of Tables on the Chemical List Presence Keyword Detail Page

Table Name	Description	Table Columns	Column Description
Documents	The Documents table shows all data documents related to a specific chemical presence tag.	Document Title	Associated Data document name, click the link to view the Data Document page.
		Tags	All chemical presence tags that appear in the associated data document at the chemical level.
Chemicals	The Chemicals table shows all chemicals found within data documents related to a specific functional use category.	DTXSID	DSSTox substance identifier (DTXSID) used on the EPA CompTox Chemicals Dashboard, links to the Chemical Detail page.
		Curated CAS	Curated CAS Registry Number Associated with DTXSID, links to the Chemical Detail page.
		Curated Chemical Name	Curated Chemical Name Associated with DTXSID.

A.5: Data Document Tables

Table A.5: Data available on the Data Document page.

Data	Data Type	Description
Data Document Title	Document Metadata	The title of the data document. e.g., if the data document is a journal article or report, this might be the title of the article or report; if the data document is an MSDS sheet, this may be the name of the product represented in the MSDS sheet. If no title is provided when the data document is registered by the curator, the title defaults to be the same as the file name of the data document.
Data Document Subtitle	Document Metadata	The document subtitle. May also contain information specific to the document table or figure extracted.
Organization	Document Metadata	Organization associated with the document. Can be the journal, manufacturer, organization that published a report, governing body, etc.
Data Source	Document Metadata	The highest-level organizational category for documents in Factotum. Documents contained within a Factotum data source may come from a single source, have the same data type, or have other characteristics in common. Data source name may refer to the organization supplying the data (e.g., Walmart) or the organization which collected the data (e.g., Danish EPA).
Data Group	Document Metadata	The second level of organization for documents. A data group belongs to only one data source, and a document may belong to only one data group. All documents within the data group belong to the same data group type.
Data Group Type	Document Metadata	An internal curation classifier used to group documents that contain the same type of information. E.g., composition, functional use, chemical presence.
Document Type	Document Metadata	Indicates the type of document format, as obtained from the original source. i.e., MSDS, journal article, governmental Report.
EPA Registration Number	Document Metadata	EPA-assigned identifier applies to documents about specific pesticides or disinfectants.
HERO ID	Document Metadata	Reference ID for EPA's Health & Environmental Research Online (HERO) Database.

Data	Data Type	Description
PMID	Document Metadata	The PubMed ID, provided if the source document is an article in the PubMed library.
Author	Document Metadata	The name of the author of an article, report, etc., as obtained from the original source.
DOI	Document Metadata	Digital Object Identifier (DOI) used to permanently identify an article or document and provide an external link.
Document URL	Document Metadata	External URL that links to the original document or page the document was downloaded from. URLs may degrade over time but were active at the time the data was originally accessed.
Extracted Text	Document Metadata	Refers to metadata extracted from the source document that is not chemical specific.
Product Name	Document Metadata	Name of data record associated with a Chemical Composition document that contains product-relevant information such as manufacturer, brand, UPC, size, item description, product images, etc. There may be multiple products associated with a single composition document, as products are based on unique Universal Product Codes (UPC).
Document Date	Document Metadata	The date the document was created or last updated, as described on the original source document.
Revision Number	Document Metadata	The Revision or Version number as reported on original source.
Reported Chemical Name	Chemical	Chemical name as provided on original source document.
Reported CAS	Chemical	Chemical CAS Registry Number (CASRN) as provided on original source document.
Curated Chemical Name	Chemical	Curated Chemical Name associated with DTXSID, provided by DSSTOX. The curated chemical name may be different than the reported chemical name.
Curated CAS	Chemical	Curated Chemical CAS Registry Number (CASRN) associated with DTXSID, provided by DSSTOX. The curated CAS may be different than the reported CAS.
DTXSID	Chemical	DSSTox Substance Identifier. Unique substance identifier, provided by DSSTOX.
Provisional Flag	Chemical	Indicates whether the chemical record was provisionally assigned a DTXSID chemical identifier [Yes/No].

Data	Data Type	Description
Raw Composition	Chemical Composition	Raw weight fraction/composition data, as reported on the source document. Unit type may be reported as percent, decimal fraction, or other form.
Cleaned Weight Fraction	Chemical Composition	Cleaned weight fraction data. Raw Composition data is curated to a harmonized form (decimal fraction between 0-1).
Ingredient Rank	Chemical Composition	The numerical order the chemical ingredient appears in an ingredient list.
Reported Function Category	Chemical Function	Functional use of the chemical in a product, as reported on the original source document.
Harmonized Function Category	Chemical Function	Harmonized identifiers describing the function a specific chemical serves (for example: fragrance, cleaning agent, filler, etc.) in a product. Function Categories are obtained from OECD or developed by EPA, and are assigned via manual curation.
Keywords Applied	Chemical Presence	This section lists all chemical presence tags associated with any chemicals that appear in the document.
Chemical Presence Keyword	Chemical Presence	Descriptive keywords assigned to chemical records, that are used to convey information about general chemical use.
Detected Flag	Chemical Presence	Indicates whether a chemical was measured or identified in environmental media or consumer products. "Yes" means that the chemical was detected by the study, while "No" could mean that the chemical was tested for, but not found; below the limit of detection for the instrumentation used; or instances where a measurement was not reported due to issues with measuring equipment, QA concerns, or other reason.
Component	Composition or Chemical Presence	Used to separate reported chemicals based on components, as reported in the source document. For example, by product component, for products with multiple parts or items in a set.

Appendix B: Download Terminology

B.1: Metadata Download Fields

B.1.1: PUC

Table B.1: PUC Metadata Download

Column Name	Description
PUC ID	The ID number used to reference the PUC in the database.
PUC General Category	Highest level in PUC Hierarchy. This is the most General Category describing a product's use (i.e., personal care).
PUC Product Family	Second Level in PUC Hierarchy. The Product Family describes a more specific type of product within the general category (i.e., dental care, body hygiene, make-up and related).
PUC Product Type	Third (most specific) Level in PUC Hierarchy. The Product Type contains a specific type of product within a Product Family (i.e., toothpaste, mouthwash).
Allowed Attributes	Attribute tag that may be applied to a product categorized into a certain PUC, based on the characteristics of the specific product (e.g. product form, demographic, micro-environment).
Assumed Attributes	Attribute tag that is automatically applied to any product categorized into a certain PUC, based on general characteristics of the PUC.
Definition	Description of the Product Use Category, including example products and exclusions.
PUC Kind	Label for Product Use Categories based on the regulatory guidelines regarding the manufacture and ingredient reporting requirements for products (FO: Formulation, AR: Article, OC: Industrial/Occupational).
PUC Level	Numerical Level of the PUC in the PUC Hierarchy [1, 2, or 3].
Product Count	The number of products that have been assigned to this specific Product Use Category.
Cumulative Product Count	The number of products within the PUC, including all lower-level PUCs within the hierarchy.

B.1.2: PUC Attribute

Table B.2: PUC Attribute Metadata Download

Column Name	Description
PUC Attribute	tags describing qualities of a particular product, such as the product form (i.e., aerosol, liquid, gel), intended user demographic (i.e., child), special features of the product (i.e., two-component), or micro-environment in which the product is used (i.e., Exterior, Interior).
Definition	Definition of PUC attribute, including example uses and exclusions.

B.1.3: List Presence Keywords

Table B.3: List Presence Keywords Metadata Download

Column Name	Description
List Presence Keyword	CPCat Chemical List Presence Keyword Name, assigned to chemical records to convey information about general chemical use.
Definition	Definition of list presence tag, which reflects how the keyword is used, or contains relevant information about the data source.
Keyword Kind	Organizational system that identifies broader categories that group related keywords (e.g., Location, PUC, Media, etc.).

B.1.4: FC

Table B.4: FC Metadata Download

Column Name	Description
Function Category	Function Category Name, obtained from OECD or developed by EPA. Harmonized identifier describing the function a specific chemical serves (e.g., fragrance, cleaning agent, filler, etc.).
Definition	Definition of Function Category, including synonyms, closely related terms and exclusions, obtained from OECD or developed by EPA.

B.2: Bulk Data Downloads

B.2.1: List Presence Keywords

Table B.5: List Presence Bulk Data Download

Column Name	Description
Data Source	The highest-level organizational category which groups data documents that may come from a single source, have the same data type, or have other characteristics in common. Data source name may refer to the organization supplying the data (e.g., Walmart) or the organization which collected the data (e.g., Danish EPA).
Data Document Title	The title of the data document. May be the title of the article or report, or the name of the product represented in the MSDS sheet. If no title is provided when the data document is registered by the curator, the title defaults to be the same as the file name of the data document.
Data Document Subtitle	The document subtitle. May also contain information specific to the document table or figure extracted.
Document Date	The date the document was created or last updated, as described on the original source document.
Organization	Organization associated with the document. Can be the journal, manufacturer, organization that published a report, governing body, etc.

Column Name	Description
Raw Chemical Name	Chemical name as provided on original source document.
Raw CAS	Chemical CAS Registry Number (CASRN) as provided on original source document.
DTXSID	DSSTox Substance Identifier. Unique substance identifier, provided by DSSTOX.
Curated Chemical Name	Curated Chemical Name associated with DTXSID, provided by DSSTOX. The curated chemical name may be different than the reported chemical name.
Curated CAS	Curated Chemical CAS Registry Number (CASRN) associated with DTXSID, provided by DSSTOX. The curated CAS may be different than the reported CAS.
Provisional Flag	Indicates whether the chemical record was provisionally assigned a DTXSID chemical identifier [Yes/No].
Reported Function Category	A reported functional use assigned to the chemical.
Harmonized Function Category	The curator assigned functional use category for the reported text.
Keyword Set	Refers to the group of chemical presence keywords assigned to a given chemical that appears in a data document. Descriptive keyword sets are used to convey information about general chemical use, and should be interpreted as a whole.

B.2.2: FC

Table B.6: FC Bulk Data Download

Column Name	Description
Data Source	The highest-level organizational category which groups data documents that may come from a single source, have the same data type, or have other characteristics in common. Data source name may refer to the organization supplying the data (e.g., Walmart) or the organization which collected the data (e.g., Danish EPA).
Data Group Type	An internal curation classifier used to group documents that contain the same type of information [composition, functional use, chemical presence].
Data Document Title	The title of the data document. May be the title of the article or report, or the name of the product represented in the MSDS sheet. If no title is provided when the data document is registered by the curator, the title defaults to be the same as the file name of the data document.

Column Name	Description
Document Date	The date the document was created or last updated, as described on the original source document.
Raw Chemical Name	Chemical name as provided on original source document.
Raw CAS	Chemical CAS Registry Number (CASRN) as provided on original source document.
DTXSID	DSSTox Substance Identifier. Unique substance identifier, provided by DSSTOX.
Curated Chemical Name	Curated Chemical Name associated with DTXSID, provided by DSSTOX. The curated chemical name may be different than the reported chemical name.
Curated CAS	Curated Chemical CAS Registry Number (CASRN) associated with DTXSID, provided by DSSTOX. The curated CAS may be different than the reported CAS.
Provisional Flag	Indicates whether the chemical record was provisionally assigned a DTXSID chemical identifier [Yes/No].
Reported Function Category	Functional use of the chemical in a product, as reported on the original source document.
Harmonized Function Category	Harmonized identifiers describing the function a specific chemical serves (for example: fragrance, cleaning agent, filler, etc). in a product. Functional use categories are obtained from OECD or developed by EPA.

B.2.3: Product Composition

Table B.7: Product Composition Bulk Data Download

Column Name	Description
Data Source	The highest-level organizational category which groups data documents that may come from a single source, have the same data type, or have other characteristics in common. Data source name may refer to the organization supplying the data (e.g., Walmart) or the organization which collected the data (e.g., Danish EPA).
Data Document Title	The title of the data document. May be the title of the article or report, or the name of the product represented in the MSDS sheet. If no title is provided when the data document is registered by the curator, the title defaults to be the same as the file name of the data document.
Data Document Subtitle	The document subtitle. May also contain information specific to the document table or figure extracted.
Document Date	The date the document was created or last updated, as described on the original source document.

Column Name	Description
Product Name	Name of data record associated with a Chemical Composition document that contains product-relevant information such as manufacturer, brand, UPC, size, item description, product images, etc. There may be multiple products associated with a single composition document, as products are based on unique Universal Product Codes (UPC).
PUC Kind	Label for Product Use Categories based on the regulatory guidelines regarding the manufacture and ingredient reporting requirements for products [FO: Formulation, AR: Article, OC: Industrial/Occupational].
PUC General Category	Highest level in PUC Hierarchy. This is the most General Category describing a product's use (i.e., personal care).
PUC Product Family	Second Level in PUC Hierarchy. The Product Family describes a more specific type of product within the general category (i.e., dental care, body hygiene, make-up and related).
PUC Product Type	Third (most specific) Level in PUC Hierarchy. The Product Type contains a specific type of product within a Product Family (i.e., toothpaste, mouthwash).
PUC Classification Method	Curation method, whether manual or automated (natural language processing-based), used to assign product to PUC (Product Use Category) [Manual, Automatic Conversion, Manual Batch, Bulk Assign, Automated].
Raw Chemical Name	Chemical name as provided on original source document.
Raw CAS	Chemical CAS Registry Number (CASRN) as provided on original source document.
DTXSID	DSSTox Substance Identifier. Unique substance identifier, provided by DSSTOX.
Curated Chemical Name	Curated Chemical Name associated with DTXSID, provided by DSSTOX. The curated chemical name may be different than the reported chemical name.
Curated CAS	Curated Chemical CAS Registry Number (CASRN) associated with DTXSID, provided by DSSTOX. The curated CAS may be different than the reported CAS.
Provisional Flag	Indicates whether the chemical record was provisionally assigned a DTXSID chemical identifier [Yes/No].
Raw Min Comp	The raw/reported minimum composition information, as reported on the original source. This is the lower limit of chemical concentration specified by a range. May be reported as percent, decimal fraction, or other form.
Raw Max Comp	The raw/reported maximum composition information, as reported on the original source. This is the upper limit of chemical concentration

Column Name	Description
	specified by a range. May be reported as percent, decimal fraction, or other form.
Raw Central Comp	The raw/reported central composition information, as reported on the original source. This is the chemical concentration if a central or single point value is reported. May be reported as percent, decimal fraction, or other form.
Unit Type	The unit type for the concentration, as reported in source document.
Lower Weight Fraction	The minimum value for cleaned chemical weight fraction data. This is reported composition data that has been cleaned to a harmonized form [decimal fraction between 0-1].
Upper Weight Fraction	The maximum value for cleaned chemical weight fraction data. This is reported composition data that has been cleaned to a harmonized form [decimal fraction between 0-1].
Central Weight Fraction	The central or point value for cleaned chemical weight fraction data. This is reported composition data that has been cleaned to a harmonized form [decimal fraction between 0-1].
Weight Fraction Type	Reported (Default) refers to weight fractions calculated using composition data present on the source document, while Predicted values are generated via model [predicted or reported].

B.3: Details Page Downloads

B.3.1: Chemical

Product Composition

Table B.8: Download Products and Weight Fractions by Chemical; this download table can be found in the top right corner of a Chemical Detail Page

Column Name	Description
Data Document Title	The title of the data document. May be the title of the article or report, or the name of the product represented in the MSDS sheet. If no title is provided when the data document is registered by the curator, the title defaults to be the same as the file name of the data document.
Document Date	The date the document was created or last updated, as described on the original source document.
Product Name	Name of data record associated with a Chemical Composition document that contains product-relevant information such as manufacturer, brand, UPC, size, item description, product images, etc. There may be multiple products associated with a single composition document, as products are based on unique Universal Product Codes (UPC).
PUC General Category	Highest level in PUC Hierarchy. This is the most General Category describing a product's use. (i.e., personal care).

Column Name	Description
PUC Product Family	Second Level in PUC Hierarchy. The Product Family describes a more specific type of product within the general category. (i.e., dental care, body hygiene, make-up and related).
PUC Product Type	Third (most specific) Level in PUC Hierarchy. The Product Type contains a specific type of product within a Product Family. (i.e., toothpaste, mouthwash).
Classification Method	Curation method, whether manual or automated (natural language processing-based), used to assign product to PUC (Product Use Category) [Manual, Automatic Conversion, Manual Batch, Bulk Assign, Automated].
Raw Min Comp	The raw/reported minimum composition information, as reported on the original source. This is the lower limit of chemical concentration specified by a range. May be reported as percent, decimal fraction, or other form.
Raw Max Comp	The raw/reported maximum composition information, as reported on the original source. This is the upper limit of chemical concentration specified by a range. May be reported as percent, decimal fraction, or other form.
Raw Central Comp	The raw/reported central composition information, as reported on the original source. This is the chemical concentration if a central or single point value is reported. May be reported as percent, decimal fraction, or other form.
Unit Type	The unit type for the concentration, as reported in source document.
Lower Weight Fraction	The minimum value for cleaned chemical weight fraction data. This is reported composition data that has been cleaned to a harmonized form [decimal fraction between 0-1].
Upper Weight Fraction	The maximum value for cleaned chemical weight fraction data. This is reported composition data that has been cleaned to a harmonized form [decimal fraction between 0-1].
Central Weight Fraction	The central or point value for cleaned chemical weight fraction data. This is reported composition data that has been cleaned to a harmonized form [decimal fraction between 0-1].
Weight Fraction Type	Reported (Default) refers to weight fractions calculated using composition data present on the source document, while Predicted values are generated via model [predicted or reported].

FC

Table B.9: Download FC by Chemical; this download table can be found in the top right corner of a Chemical Detail Page

Column Name	Description
Data Group Type	An internal curation classifier used to group documents that contain the same type of information [composition, functional use, chemical presence].
Data Document Title	The title of the data document. May be the title of the article or report, or the name of the product represented in the MSDS sheet. If no title is provided when the data document is registered by the curator, the title defaults to be the same as the file name of the data document.
Document Date	The date the document was created or last updated, as described on the original source document.
Reported Function Category	Functional use of the chemical in a product, as reported on the original source document.
Harmonized Function Category	Harmonized identifiers describing the function a specific chemical serves (for example: fragrance, cleaning agent, filler, etc.) in a product. Functional use categories are obtained from OECD or developed by EPA.

B.3.2: PUC

Product Composition

Table B.10: Download Products and Chemical Weight Fractions by PUC; this download table can be found the top of the Information Categories tables on a PUC Detail Page

Column Name	Description
Data Document Title	The title of the data document. May be the title of the article or report, or the name of the product represented in the MSDS sheet. If no title is provided when the data document is registered by the curator, the title defaults to be the same as the file name of the data document.
Document Date	The date the document was created or last updated, as described on the original source document.
Product Name	Name of data record associated with a Chemical Composition document that contains product-relevant information such as manufacturer, brand, UPC, size, item description, product images, etc. There may be multiple products associated with a single composition document, as products are based on unique Universal Product Codes (UPC).
PUC Classification Method	Curation method, whether manual or automated (natural language processing-based), used to assign product to PUC (Product Use

Column Name	Description
	Category) [Manual, Automatic Conversion, Manual Batch, Bulk Assign, Automated].
Raw Chemical Name	Chemical name as provided on original source document.
Raw CAS Number	Chemical CAS Registry Number (CASRN) as provided on original source document.
DTXSID	DSSTox Substance Identifier. Unique substance identifier, provided by DSSTOX.
Curated Chemical Name	Curated Chemical Name associated with DTXSID, provided by DSSTOX. The curated chemical name may be different than the reported chemical name.
Curated CAS Number	Curated Chemical CAS Registry Number (CASRN) associated with DTXSID, provided by DSSTOX. The curated CAS may be different than the reported CAS.
Raw Min Comp	The raw/reported minimum composition information, as reported on the original source. This is the lower limit of chemical concentration specified by a range. May be reported as percent, decimal fraction, or other form.
Raw Max Comp	The raw/reported maximum composition information, as reported on the original source. This is the upper limit of chemical concentration specified by a range. May be reported as percent, decimal fraction, or other form.
Raw Central Comp	The raw/reported central composition information, as reported on the original source. This is the chemical concentration if a central or single point value is reported. May be reported as percent, decimal fraction, or other form.
Unit Type	The unit type for the concentration, as reported in source document.
Lower Weight Fraction	The minimum value for cleaned chemical weight fraction data. This is reported composition data that has been cleaned to a harmonized form [decimal fraction between 0-1].
Upper Weight Fraction	The maximum value for cleaned chemical weight fraction data. This is reported composition data that has been cleaned to a harmonized form [decimal fraction between 0-1].
Central Weight Fraction	The central or point value for cleaned chemical weight fraction data. This is reported composition data that has been cleaned to a harmonized form [decimal fraction between 0-1].
Weight Fraction Type	Reported (Default) refers to weight fractions calculated using composition data present on the source document, while Predicted values are generated via model [predicted or reported].

Chemical

Table B.11: Download Chemicals by PUC; this download table can be found the top of the Information Categories tables on a PUC Detail Page

Column Name	Description
DTXSID	DSSTox Substance Identifier. Unique substance identifier, provided by DSSTOX
Curated CAS	Curated Chemical Name associated with DTXSID, provided by DSSTOX. The curated chemical name may be different than the reported chemical name.
Curated Chemical Name	Curated Chemical CAS Registry Number (CASRN) associated with DTXSID, provided by DSSTOX. The curated CAS may be different than the reported CAS.
Manual Count	The number of times the chemical appears in data documents associated with the PUC of interest, where the products have been curated via manual classification methods.
Automatic Count	The number of times the chemical appears in data documents associated with the PUC of interest, where the products have been curated via the Automatic (natural language processing-based) classification method.

FC

Table B.12: Download FC by PUC; this download table can be found the top of the Information Categories tables on a PUC Detail Page

Column Name	Description
Data Document Title	The title of the data document. May be the title of the article or report, or the name of the product represented in the MSDS sheet. If no title is provided when the data document is registered by the curator, the title defaults to be the same as the file name of the data document.
Raw Chemical Name	Chemical name as provided on original source document.
Raw CAS	Chemical CAS Registry Number (CASRN) as provided on original source document.
DTXSID	DSSTox Substance Identifier. Unique substance identifier, provided by DSSTOX.
Curated Chemical Name	Curated Chemical Name associated with DTXSID, provided by DSSTOX. The curated chemical name may be different than the reported chemical name.
Curated CAS	Curated Chemical CAS Registry Number (CASRN) associated with DTXSID, provided by DSSTOX. The curated CAS may be different than the reported CAS.
Reported Functional Use	Functional use of the chemical in a product, as reported on the original source document.

Column Name	Description
Harmonized Functional Use	Harmonized identifiers describing the function a specific chemical serves (for example: fragrance, cleaning agent, filler, etc.) in a product. Functional use categories are obtained from OECD or developed by EPA.

B.3.3: FC

Product and PUC

Table B.13: Download Products by FC

Column Name	Description
Data Document Title	The title of the data document. May be the title of the article or report, or the name of the product represented in the MSDS sheet. If no title is provided when the data document is registered by the curator, the title defaults to be the same as the file name of the data document.
Document Date	The date the document was created or last updated, as described on the original source document.
Product Name	Name of data record associated with a Chemical Composition document that contains product-relevant information such as manufacturer, brand, UPC, size, item description, product images, etc. There may be multiple products associated with a single composition document, as products are based on unique Universal Product Codes (UPC).
PUC General Category	Highest level in PUC Hierarchy. This is the most General Category describing a product's use (i.e., personal care).
PUC Product Family	Second Level in PUC Hierarchy. The Product Family describes a more specific type of product within the general category (i.e., dental care, body hygiene, make-up and related).
PUC Product Type	Third (most specific) Level in PUC Hierarchy. The Product Type contains a specific type of product within a Product Family (i.e., toothpaste, mouthwash).
PUC Classification Method	Curation method, whether manual or automated (natural language processing-based), used to assign product to PUC (Product Use Category) [Manual, Automatic Conversion, Manual Batch, Bulk Assign, Automated].
Raw Chemical Name	Chemical name as provided on original source document.
Raw CAS	Chemical CAS Registry Number (CASRN) as provided on original source document.
DTXSID	DSSTox Substance Identifier. Unique substance identifier, provided by DSSTOX.

Column Name	Description
Curated Chemical Name	Curated Chemical Name associated with DTXSID, provided by DSSTOX. The curated chemical name may be different than the reported chemical name.
Curated CAS	Curated Chemical CAS Registry Number (CASRN) associated with DTXSID, provided by DSSTOX. The curated CAS may be different than the reported CAS.
Reported Function Category	Functional use of the chemical in a product, as reported on the original source document.

Chemical

Table B.14: Download Chemicals by FC

Column Name	Description
DTXSID	DSSTox Substance Identifier. Unique substance identifier, provided by DSSTOX.
Curated Chemical Name	Curated Chemical Name associated with DTXSID, provided by DSSTOX. The curated chemical name may be different than the reported chemical name.
Curated CAS	Curated Chemical CAS Registry Number (CASRN) associated with DTXSID, provided by DSSTOX. The curated CAS may be different than the reported CAS.

B.3.4: Data Document

Composition

Table B.15: Download Chemical Data for a Composition Data Document

Column Name	Description
Data Source	The highest-level organizational category which groups data documents that may come from a single source, have the same data type, or have other characteristics in common. Data source name may refer to the organization supplying the data (e.g., Walmart) or the organization which collected the data (e.g., Danish EPA).
Data Document Title	The title of the data document. May be the title of the article or report, or the name of the product represented in the MSDS sheet. If no title is provided when the data document is registered by the curator, the title defaults to be the same as the file name of the data document.
Data Document Subtitle	The document subtitle. May also contain information specific to the document table or figure extracted.
Document Date	The date the document was created or last updated, as described on the original source document.

Column Name	Description
Product Name	Name of data record associated with a Chemical Composition document that contains product-relevant information such as manufacturer, brand, UPC, size, item description, product images, etc. There may be multiple products associated with a single composition document, as products are based on unique Universal Product Codes (UPC).
Raw Chemical Name	Chemical name as provided on original source document.
Raw CAS	Chemical CAS Registry Number (CASRN) as provided on original source document.
DTXSID	DSSTox Substance Identifier. Unique substance identifier, provided by DSSTOX.
Curated Chemical Name	Curated Chemical Name associated with DTXSID, provided by DSSTOX. The curated chemical name may be different than the reported chemical name.
Curated CAS	Curated Chemical CAS Registry Number (CASRN) associated with DTXSID, provided by DSSTOX. The curated CAS may be different than the reported CAS.
Provisional Flag	Indicates whether the chemical record was provisionally assigned a DTXSID chemical identifier [Yes/No].
Has Composition Data	Flag indicating whether the chemical record contains reported composition data [Yes/No].
Raw Min Comp	The raw/reported minimum composition information, as reported on the original source. This is the lower limit of chemical concentration specified by a range. May be reported as percent, decimal fraction, or other form.
Raw Max Comp	The raw/reported maximum composition information, as reported on the original source. This is the upper limit of chemical concentration specified by a range. May be reported as percent, decimal fraction, or other form.
Raw Central Comp	The raw/reported central composition information, as reported on the original source. This is the chemical concentration if a central or single point value is reported. May be reported as percent, decimal fraction, or other form.
Unit Type	The unit type for the concentration, as reported in source document.
Lower Weight Fraction	The minimum value for cleaned chemical weight fraction data. This is reported composition data that has been cleaned to a harmonized form [decimal fraction between 0-1].

Column Name	Description
Upper Weight Fraction	The maximum value for cleaned chemical weight fraction data. This is reported composition data that has been cleaned to a harmonized form [decimal fraction between 0-1].
Central Weight Fraction	The central or point value for cleaned chemical weight fraction data. This is reported composition data that has been cleaned to a harmonized form [decimal fraction between 0-1].
Weight Fraction Type	Reported (Default) refers to weight fractions calculated using composition data present on the source document, while Predicted values are generated via model [predicted or reported].
Ingredient Rank	The numerical order the chemical ingredient appears in an ingredient list.
Component	Used to separate reported chemicals based on components, as reported in the source document. For example, by product component, for products with multiple parts or items in a set.
Reported Function Category	Functional use of the chemical in a product, as reported on the original source document.
Harmonized Function Category	Harmonized identifiers describing the function a specific chemical serves (for example: fragrance, cleaning agent, filler, etc.) in a product. Functional use categories are obtained from OECD or developed by EPA.

List Keyword Presence

Table B.16: Download Chemical Data for a Chemical Presence Data Document

Column Name	Description
Data Document Title	The title of the data document. May be the title of the article or report, or the name of the product represented in the MSDS sheet. If no title is provided when the data document is registered by the curator, the title defaults to be the same as the file name of the data document.
Data Document Subtitle	The document subtitle. May also contain information specific to the document table or figure extracted.
Organization	Organization associated with the document. Can be the journal, manufacturer, organization that published a report, governing body, etc.
Document Date	The date the document was created or last updated, as described on the original source document.
Raw chemical name	Chemical name as provided on original source document.
Raw CAS	Chemical CAS Registry Number (CASRN) as provided on original source document.

Column Name	Description
DTXSID	DSSTox Substance Identifier. Unique substance identifier, provided by DSSTOX.
Curated Chemical Name	Curated Chemical Name associated with DTXSID, provided by DSSTOX. The curated chemical name may be different than the reported chemical name.
Curated CAS	Curated Chemical CAS Registry Number (CASRN) associated with DTXSID, provided by DSSTOX. The curated CAS may be different than the reported CAS.
Provisional Flag	Indicates whether the chemical record was provisionally assigned a DTXSID chemical identifier [Yes/No].
Reported Function Category	Functional use of the chemical in a product, as reported on the original source document.
Harmonized Function Category	Harmonized identifiers describing the function a specific chemical serves (for example: fragrance, cleaning agent, filler, etc. in a product. Functional use categories are obtained from OECD or developed by EPA.
Keyword Set	Refers to the group of chemical presence keywords assigned to a given chemical that appears in a data document. Descriptive keyword sets are used to convey information about general chemical use, and should be interpreted as a whole.
Component	Used to separate reported chemicals based on components, as reported in the source document. For example, by product component, for products with multiple parts or items in a set.

Function Category

Table B.17: Download Chemical Data for a Functional Use Data Document

Column Name	Description
Data Document Title	The title of the data document. May be the title of the article or report, or the name of the product represented in the MSDS sheet. If no title is provided when the data document is registered by the curator, the title defaults to be the same as the file name of the data document.
Data Document Subtitle	The document subtitle. May also contain information specific to the document table or figure extracted.
Organization	Organization associated with the document. Can be the journal, manufacturer, organization that published a report, governing body, etc.
Raw chemical name	Chemical name as provided on original source document.
Raw CAS	Chemical CAS Registry Number (CASRN) as provided on original source document.

Column Name	Description
DTXSID	DSSTox Substance Identifier. Unique substance identifier, provided by DSSTOX.
Curated Chemical Name	Curated Chemical Name associated with DTXSID, provided by DSSTOX. The curated chemical name may be different than the reported chemical name.
Curated CAS	Curated Chemical CAS Registry Number (CASRN) associated with DTXSID, provided by DSSTOX. The curated CAS may be different than the reported CAS.
Provisional Flag	Indicates whether the chemical record was provisionally assigned a DTXSID chemical identifier [Yes/No].
Reported Functional Use	Functional use of the chemical in a product, as reported on the original source document.
Harmonized Functional Use	Harmonized identifiers describing the function a specific chemical serves (for example: fragrance, cleaning agent, filler, etc.) in a product. Functional use categories are obtained from OECD or developed by EPA.

Appendix C: A Primer on Safety Data Sheets

Safety Data Sheets (SDSs) are a publication that is mandated by the U.S. Occupational Safety and Health Administration (OSHA) under the 2012 revision of Hazard Communication Standard (HCS) (29 CFR 1910.1200(g)) (U.S. Occupational Health and Safety Administration (2012)). Under the HCS, manufacturers of substances and mixtures or substances (for example, consumer products) are required to communicate the hazards of chemical in the substance or mixture to downstream users. SDSs have many different uses ranging from allowing those using chemicals in the lab to know how substances may behave and to take appropriate precautions to informing consumers of the hazards of chemicals in a purchased consumer product. While the SDS are required to have 16 different sections, the section of most use within chemical exposure modeling (and currently with ChemExpo) is Section 3: Composition/Information on Ingredients. In this section, the HCS requires manufacturers to provide the information about each component covered by the SDS:

- name, other common names and synonyms
- Chemical Abstract Service Registry Number (CASRN) and other unique identifiers
- exact quantities of that substance in the product
- any information on the impurities or stabilizing agents used in the product

Some points to call attention to here are that 1) if a substance in a product does not pose a known health hazard, it does not have to be reported on the SDS, and 2) if a product is considered an “article” no information on the SDS must be reported. For reference an article is defined in the HCS as following:

a manufactured item: (i) which is formed to a specific shape or design during manufacture; (ii) which has end use function(s) dependent in whole or in part upon its shape or design during end use; and (iii) which does not release, or otherwise result in exposure to, a hazardous chemical under normal conditions of use.

It should also be noted that while the HCS does require exact percentages of chemicals, their impurities, and any stabilizing additives be reported on SDSs, there are several exemptions available for chemicals in a mixture. Namely, manufacturers are allowed to report a range of concentration for chemicals in that mixture if the manufacturer claims there is a trade secret, there may be variations from one batch of product to the next, or if the SDS is used for products which are very similar in composition.

Due to these exceptions and exemptions available for both mixtures and articles under the HCS, users may note that there are products in ChemExpo for which there are no chemical cards. This is either typically because there were no reported chemicals on the SDS because it was a consumable product and no ingredients were considered a hazard to human health or because the product is considered an “article” and requires no reporting.

Further, users of ChemExpo may notices that some chemical substances reported in ChemExpo from SDS have a range for a chemical’s composition rather than a single value.

This happens due to one of the exemptions to reporting the exact chemical concentration for chemical mixtures.

For more information about what an SDS is and what information is contained within the SDS, see the [OSHA Brief on Safety Data Sheets](#).