The Chemical and Product Database and the ChemExpo Knowledgebase

Kristin Isaacs

CCTE NAMs Training Workshop

Research Triangle Park, NC

April 24 – 25, 2024

Disclaimer: The views presented are those of the author and do not represent policies or views of the U.S. EPA.
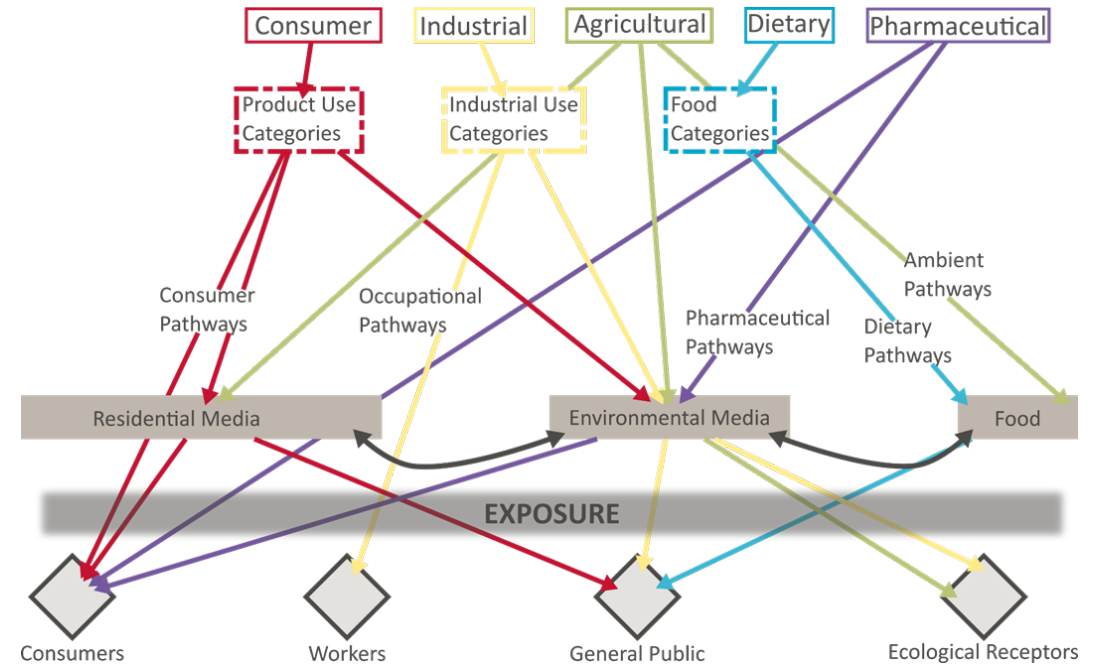
Background

- Humans interact with thousands of chemicals on a regular basis

- We often don't know precisely…
    - …which chemicals we contact in our everyday lives
    - …nor the quantities of those chemicals

- Tools and data being developed by EPA's ORD help characterize chemical use and thus exposure potential to many chemicals along various exposure pathways

- Exposure potential must explicitly be considered when prioritizing chemicals for further study, e.g., under the Toxic Substances Control Act (TSCA)
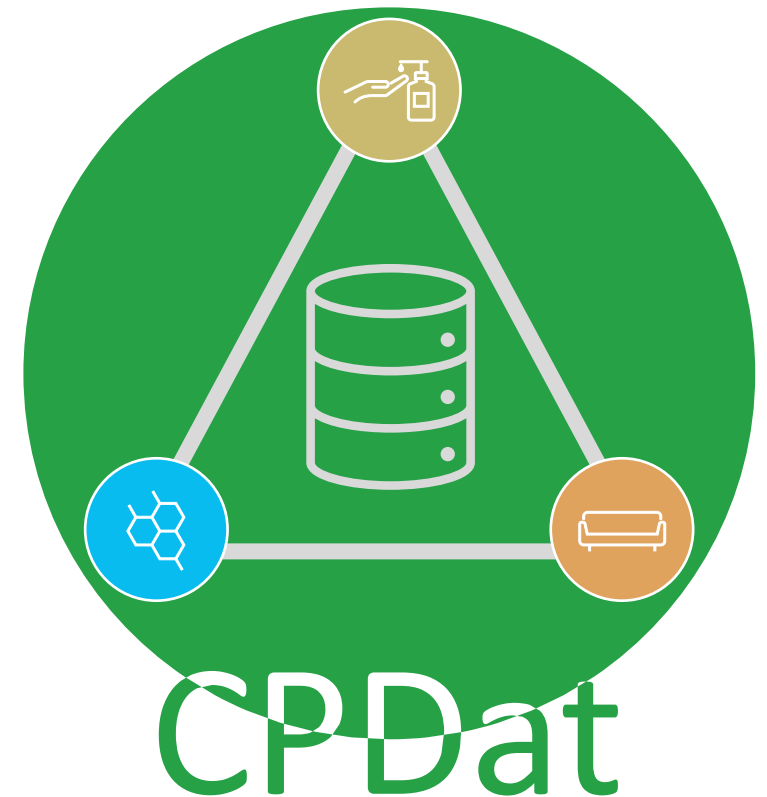
Chemical Exposure

Background

- We can often trace a chemical's exposure pathways by examining its **_use_** in various commercial sectors

- **_Chemical use_** – a primary component of chemical source – determines where and how much a chemical is released to residential or environmental compartments

- Commercial sector of use allows us to then define potential pathways associated with a chemical (potential chemical release, media occurrence, and ultimate receptor exposure)

- Use data (both qualitative and quantitative) are needed to inform models for various pathways

Exposure Pathways

Purpose of The Chemicals and Products Database

- Chemicals and Products Database (CPDat) contains information about how **chemicals are used in commerce**, including consumer products, to inform exposure assessment

- Organized by a set of defined vocabularies describing consumer product use, general commercial use, and functional use

- Data are extracted from public documents using standardized workflows

https://www.epa.gov/chemical-research/chemical-and-products-database-cpdat

CPDat

Purpose of ChemExpo Knowledgebase

- Publicly accessible, data search and visualization tool for exploring detailed data in CPDat, including key document and source metadata

- Search capability by documents' extracted text, product or product category name, chemical name, keyword, or function

- Interactive data exploration via unique landing pages for
  - product category
  - chemical
  - function category
  - Interactive data downloads

Welcome to ChemExpo Knowledgebase

ChemExpo is an interactive web application for exploring chemical data, curated from public documents, relevant to exposure assessment. ChemExpo currently surfaces data collected by EPA about how chemicals are used in commerce and how they occur in consumer and industrial products; these data are collectively known as the Chemicals and Products Database (CPDat). 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.

This beta version of ChemExpo (v0.1) is currently undergoing review and feedback from stakeholders. Its data and functionality are subject to change.

Documents ⓘ	Products ⓘ	Products Linked To Product Use Categories ⓘ
434,714	376,113	263,824

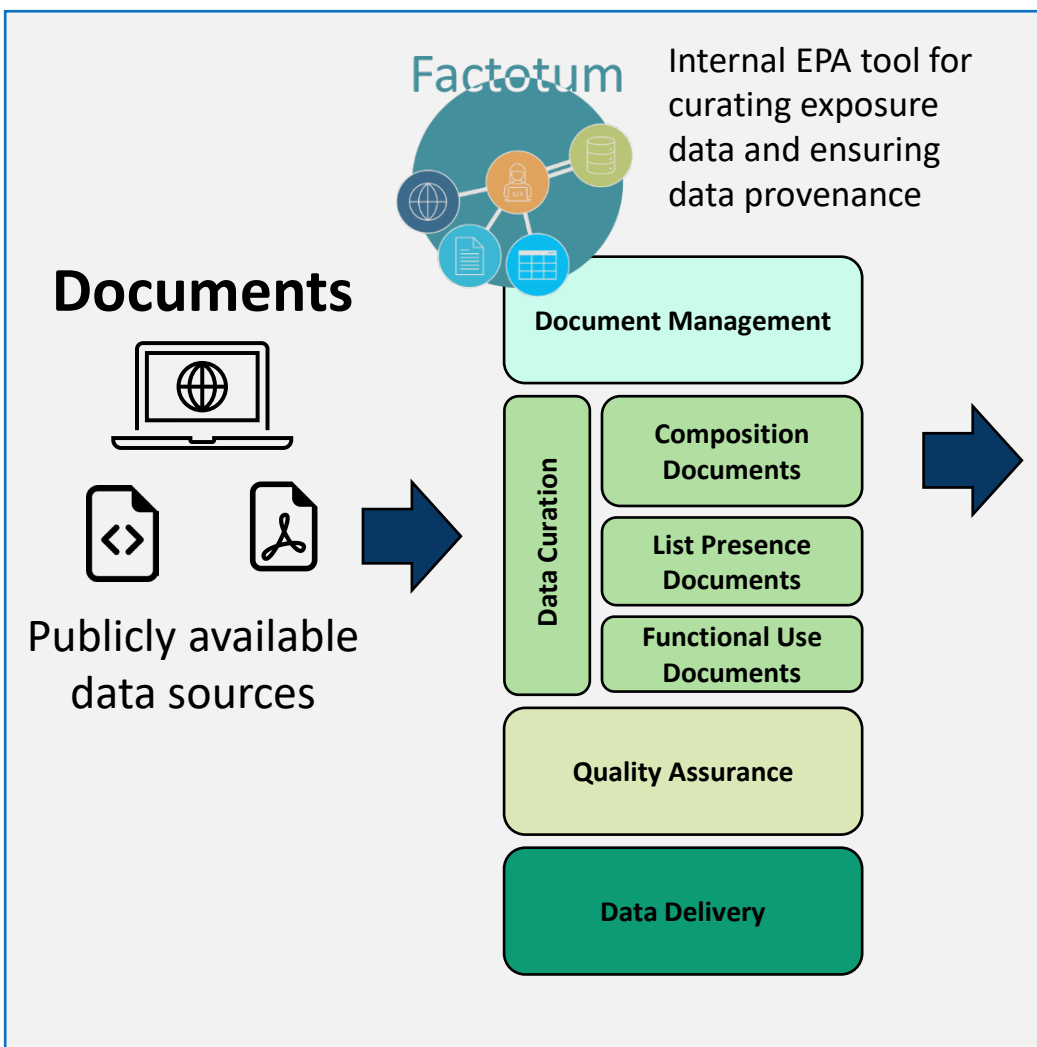| Extracted Chemical Records ⓘ	Curated Chemical Records ⓘ	Unique Chemicals (DTXSIDs) ⓘ
3.6 million	2.3 million	35,057

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
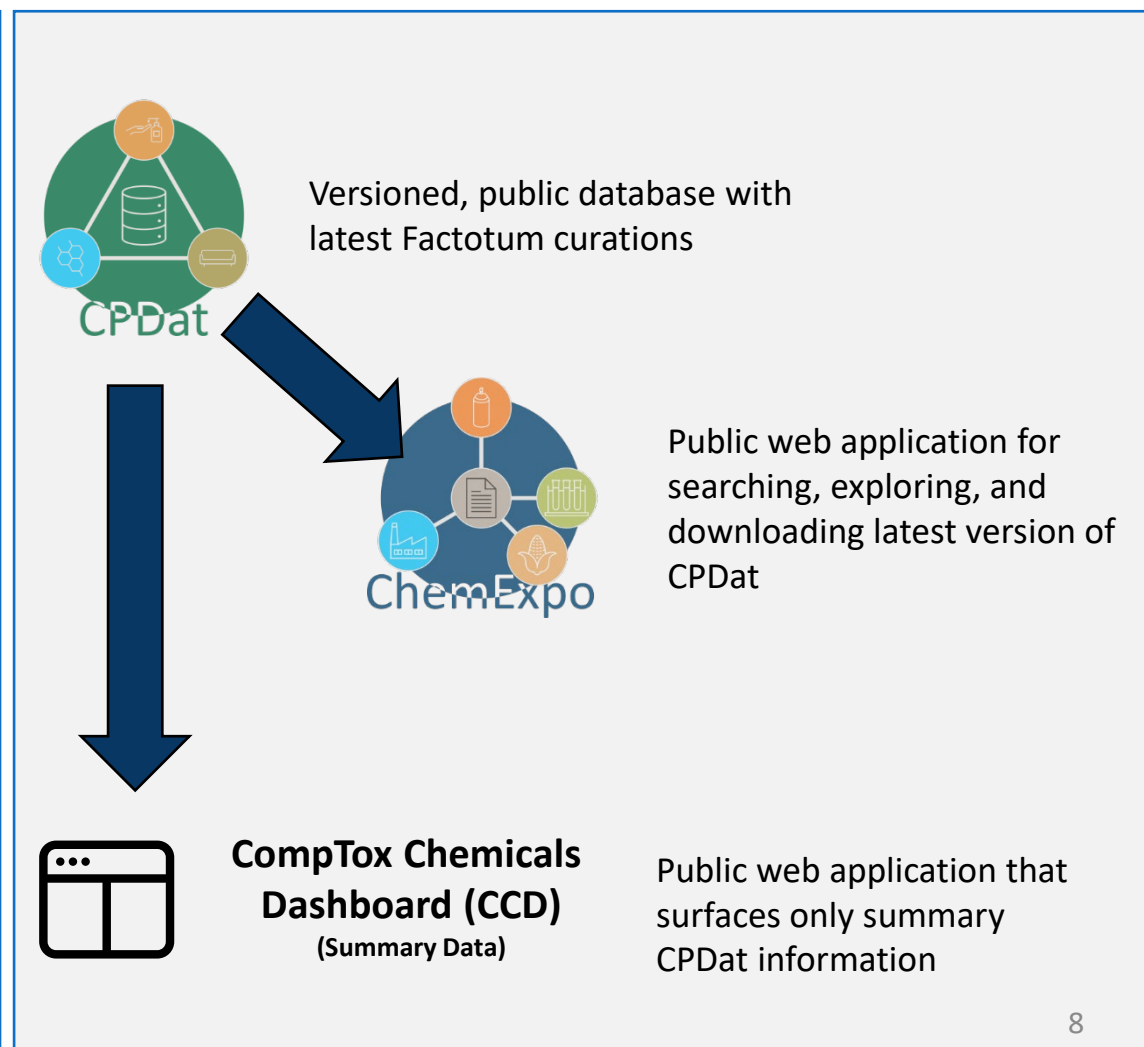
Data and Connections

- ChemExpo and CPDat make use of harmonized chemical identifiers (DTXSIDs) used in DSSTox and the CompTox Chemicals Dashboard

- CPDat summary information available via the Dashboard

- Bulk versioned downloads of each CPDat release are available

- EPA's Factotum curation tool is used for the standardized data extraction, data curation, and quality assurances workflows, ensuring for data provenance
  - Factotum is available to internal EPA users, as it may contain data and raw documents not publicly released.

Data and Connections

Internal Data Curation

Public Data Delivery

Factotum — Internal EPA tool for curating exposure data and ensuring data provenance

Documents

Publicly available data sources

Data Curation

- Document Management
- Composition Documents
- List Presence Documents
- Functional Use Documents
- Quality Assurance
- Data Delivery

CPDat — Versioned, public database with latest Factotum curations

ChemExpo — Public web application for searching, exploring, and downloading latest version of CPDat

CompTox Chemicals Dashboard (CCD) (Summary Data) — Public web application that surfaces only summary CPDat information

Example of Tool/Data Use

- CPDat is used to parameterize high-throughput models of consumer exposure (SHEDS-HT) in support of chemical prioritization.
  - Key component of EPA's ExpoCast Systematic Empirical Evaluation of Models (SEEM) consensus exposure predictions.
  - Used by the California Department of Toxic Substance Control in their chemical evaluation and priority product evaluation efforts.
  - Characterization of potential down-the-drain releases for chemicals undergoing TSCA risk evaluation.

Example of Tool/Data Use

- CPDat chemical use data have also been used to evaluate chemical use in other prioritization workflows.
    - Identification of potential children's exposures for a TSCA prioritization candidate proof-of-concept workflow and for the Toxic Free Kids program of the Minnesota Department of Health (MDH).
    - Used to characterize potential consumer exposures in releases in MDH's Contaminants of Emerging Concern (CEC program).

Considerations

- Beta public version of ChemExpo (v0.1) is currently undergoing external peer review and feedback from stakeholders.

- Its data and functionality are subject to change.

Accessing ChemExpo

- https://comptox.epa.gov/chemexpo
- https://comptox.epa.gov/chemexpo/get_data/
- APIs in development: https://api-ccte.epa.gov/docs/exposure.html

Points of Contact

Kristin Isaacs
isaacs.Kristin@epa.gov

Katherine Phillips
phillips.katherine@epa.gov

Sakshi Handa
Data Curation Lead
Handa.sakshi@epa.gov