Delivering an integrated data hub for per- and polyfluoroalkyl (PFAS) chemicals via the US EPA CompTox Chemicals Dashboard

Antony Williams[1], Chris Grulke[1], Kamel Mansouri[2], Grace Patlewicz[1] and Ann Richard[1]

[1]U.S. Environmental Protection Agency, Office of Research and Development, National Center for Computational Toxicology, Research Triangle Park, NC and [2]Integrated Laboratory Systems, Research Triangle Park, NC.

ACS Fall 2019
San Diego, CA
August 25-29, 2019

ORCID: 0000-0002-2668-4821
Antony Williams I williams.antony@epa.gov I 919-541-1033

Problem Definition and Goals

Problem: There are many sources of PFAS data online to support computational toxicology. However, curated datasets for the thousands of known PFAS chemicals are not available in structured formats,

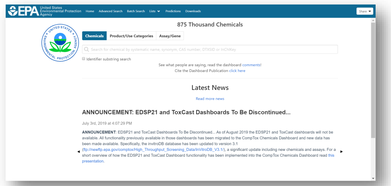
Goals: Deliver online access to hundreds of thousands of chemicals of interest to environmental science and computational toxicology. Provide lists of PFAS substances via a simple to use web-based interface. Deliver application to support diverse types of data including experimental and predicted physicochemical properties, *in vivo* hazard data and *in vitro* toxicity and toxicokinetic data. Make the data available as downloadable data for reuse and repurposing in other databases.

Abstract

There is increasing interest in the environmental impact of per- and polyfluoroalkyl substances (PFAS) chemicals and the aggregation of related data provides both agency and public access to support research. The US-EPA's CompTox Chemicals Dashboard (https://comptox.epa.gov/dashboard) is a publicly accessible website providing access to data for ~875,000 chemical substances, including thousands of PFAS chemicals. The dashboard provides access to a wide array of experimental and predicted physicochemical properties, *in vitro* bioactivity and *in vivo* toxicity data, product use information and integrated linkages to a growing list of literature, toxicology, and analytical chemistry websites. The assembly of data has required hundreds of hours of manual curation and data checking to deliver a number of segregated lists of PFAS chemicals. Experimental data extracted from literature articles has allowed the optimization of QSAR models for the prediction of properties such as logP (octanol-water partition) and aqueous solubility. This poster provides an overview of the dashboard, the ongoing expansion of the PFAS chemical library, with associated categorization, and new physicochemical property and environmental fate and transport QSAR prediction models being developed for these chemicals. Our work in the measurement and analysis of bioactivity data for PFAS chemicals will also be reviewed.

This abstract does not necessarily represent the views or policies of the U.S. Environmental Protection Agency.

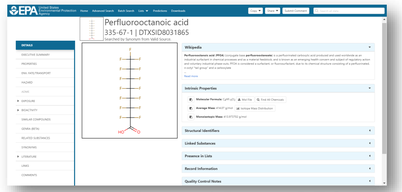
The CompTox Chemicals Dashboard



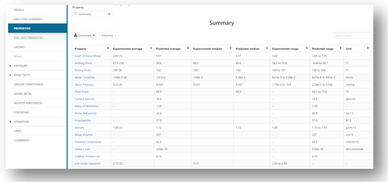
Dashboard Entry Page

The landing page of the dashboard is a simple text entry box allowing a type-ahead search for systematic, trade and trivial names, CAS Registry Numbers and InChI chemical identifiers.

Where possible, links are provided to related Wikipedia articles. Structure file formats are available for download to the desktop (SMILES and molfile) and an executive summary report regarding chemical toxicity is provided.



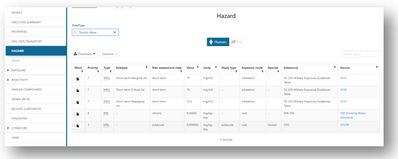
Chemical Record Page: PFOA



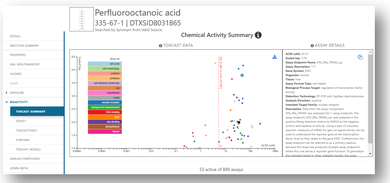
Chemical Properties Panel

For records with chemical structures experimental and predicted physicochemical (logP, water solubility etc.) and fate and transport properties are provided. These include OPERA models reported by Mansouri *et al.* [1]

The Hazard tab provides access to data assembled from a series of public resources including EPA data (i.e. IRIS and PPRTV reports, ToxRef DB). Data can be downloaded as TSV and Excel files.



Toxicity Values Panel

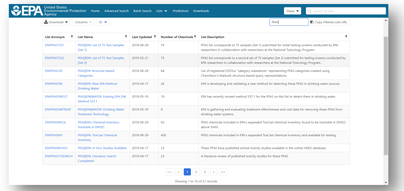


Bioactivities: e.g. ToxCast Data

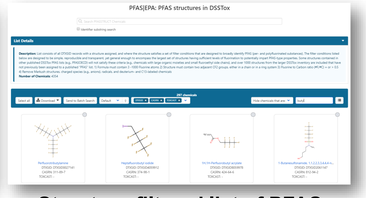
ToxCast Bioactivity data measured over the past decade are under the Bioassay Tab. Data can be downloaded as Excel files. New *in vitro* data are being generated on a library of ~150 PFAS in collaboration with NTP.

Accessing PFAS Chemical Lists

The dashboard provides access to ~20 individual PFAS chemical lists. These include a list based on structural filters: (1) Formula must contain 3 -1000 Fluorine atoms; 2) Structure must contain two adjacent CF2 groups, either in a chain or in a ring system; 3) Fluorine to Carbon ratio (#F/#C) = or > 0.5; 4) **REMOVE** charged species (e.g., bare anions), radicals, and deuterium- and C13-labeled chemicals

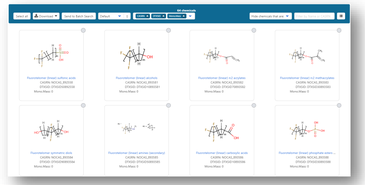


List of chemical lists of PFAS chemicals: 21 lists and growing



Structure-filtered list of PFAS

A list of Markush Representations associated with the selection of chemicals for *in vitro* toxicity and toxicokinetic screening is also available [2,3].



Future Work

- *In vitro* toxicity and toxicokinetic measurements are underway for ~150 PFAS [3]. These will be released in the future on the dashboard.
- Experimental property data are being harvested from literature and online resources to include into the dashboard and OPERA models.
- Chemical categorization efforts continue in order to be encompassing of more of the PFAS library.

References

1. Mansouri *et al.* OPERA models for predicting physchem properties and environmental fate endpoints, *J. ChemInf.* **10**, 10 (2018)
2. PFAS Categories list on the CompTox Chemicals Dashboard https://comptox.epa.gov/dashboard/chemical_lists/EPAPFASCAT
3. Patlewicz *et al.* A Chemical Category-Based Prioritization Approach for Selecting 75 Per- and Polyfluoroalkyl Substances (PFAS) for Tiered Toxicity and Toxicokinetic Testing. *Environ Health Perspect.* 2019 Jan;127(1):14501.

Acknowledgements

The authors thank the chemical curation team for their rigorous work and the software development team for the development of the dashboard.