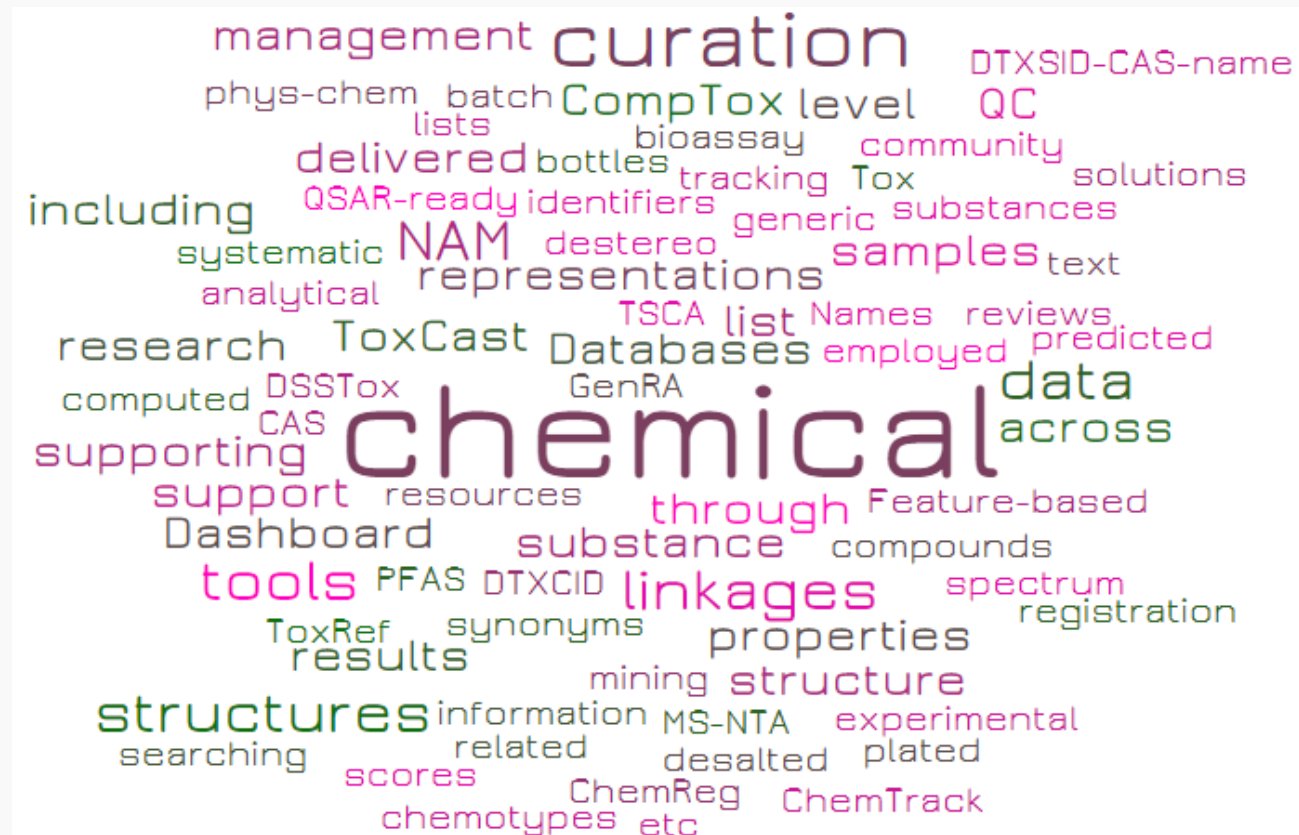


EPA's DSSTox Database: The strategic role and requirements of chemical curation

Ann M. Richard

*Center for Computational Toxicology & Exposure
Office of Research & Development
US Environmental Protection Agency*



*The views expressed in this presentation do not necessarily
represent the views or policies of the EPA.*

Linking data resources to support EPA Programs

- ToxCast (EPA) & Tox21 (Multi-Agency)
 - High-throughput screening >4000 (ToxCast) to >10K (Tox21) environmentally relevant chemicals across 10's to 100's of assays
- ACToR, ExpoCast, CPDat, ToxRef DB, TSCA, PFAS, EcoTox Lots of lists!!
 - meshed CAS lists, product-use database, *in vivo* reference data, etc.
- *In Silico* prediction models
 - exposure, ER/AR binding, phys-chem properties, ADME
- Public-facing CompTox Chemicals Dashboard, <https://ccte-cdd.epa.gov/dashboard/>
 - link data resources, facilitate access to & utility of EPA and public data

*Chemical
databases*

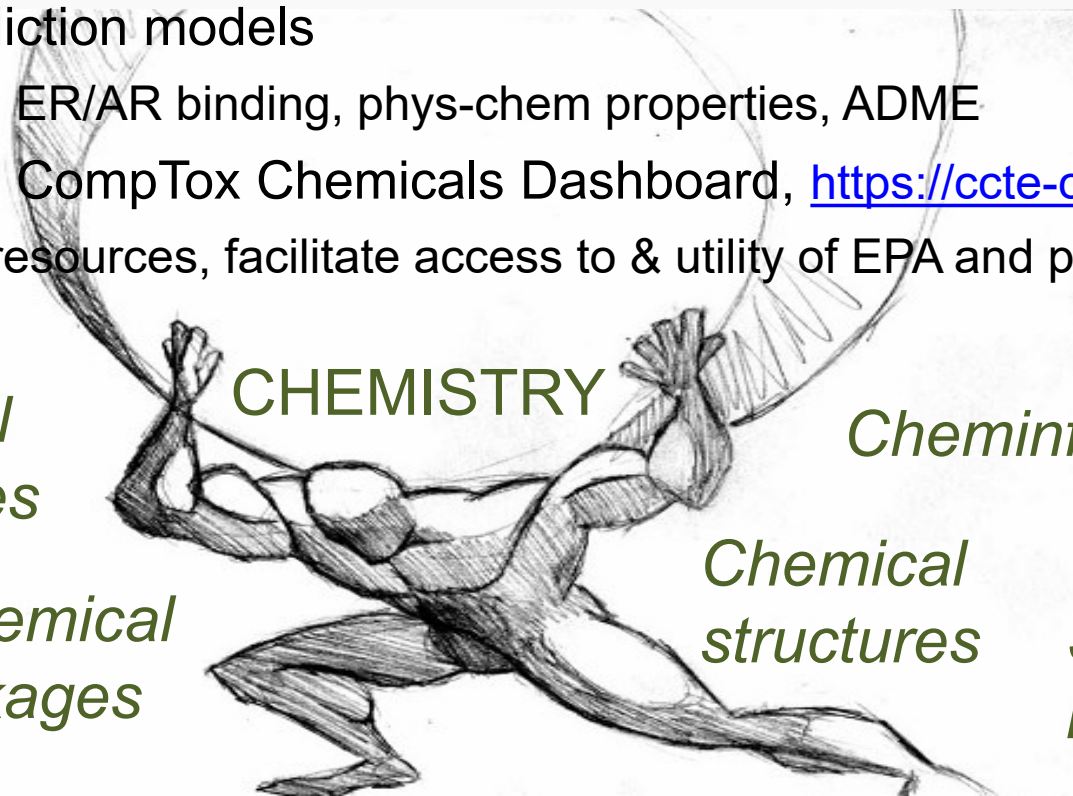
CHEMISTRY

Cheminformatics

*Chemical
linkages*

*Chemical
structures*

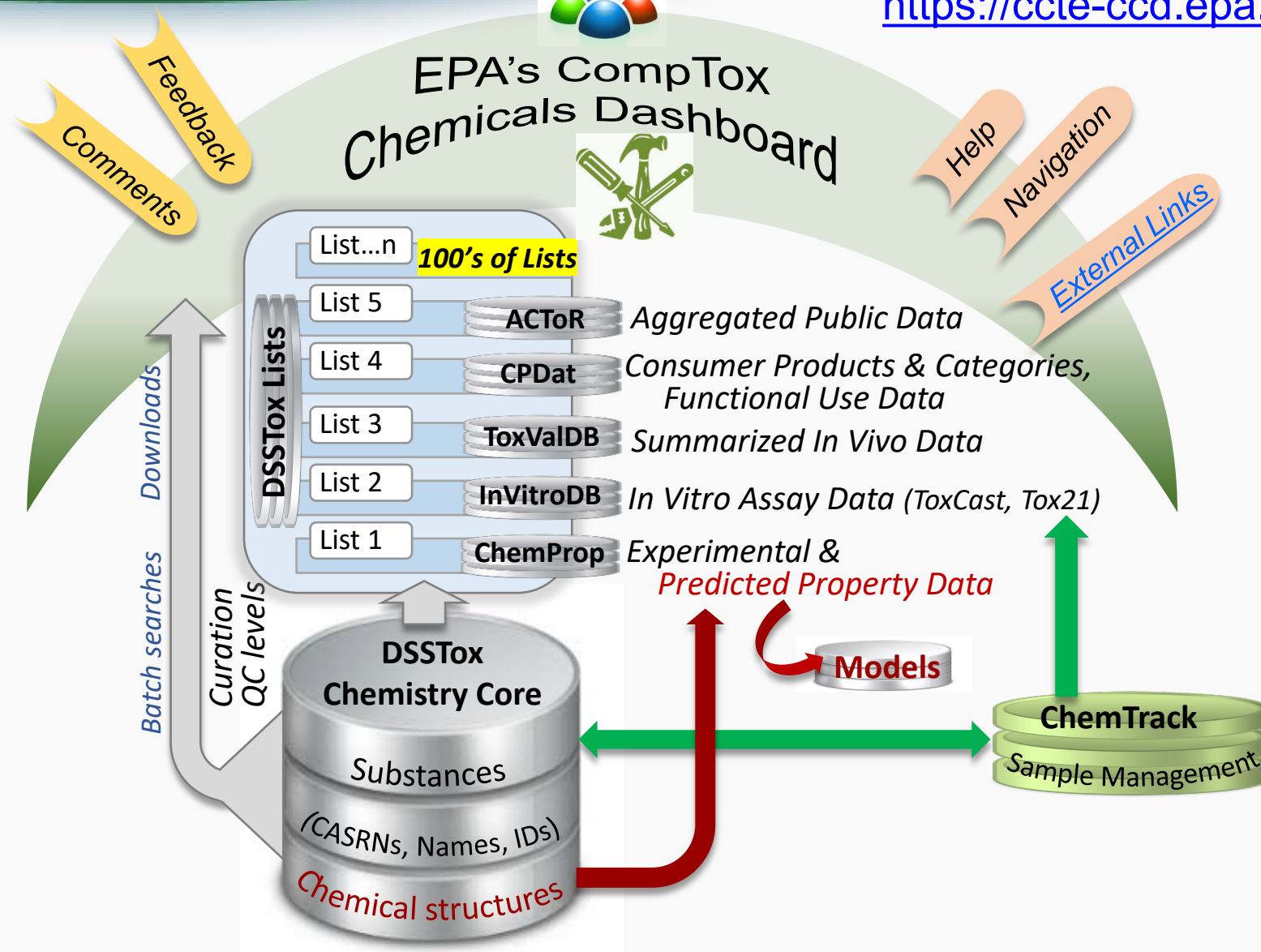
*SAR/QSAR
models*



The Chemistry beast feeding the CompTox Chemicals Dashboard!



<https://ccte-ccd.epa.gov/dashboard/>



Why build yet another public chemistry database and web-interface?

- Focus on environmental chemicals
- Focus on quality data-structure associations
- Better support EPA research & chemical regulatory programs

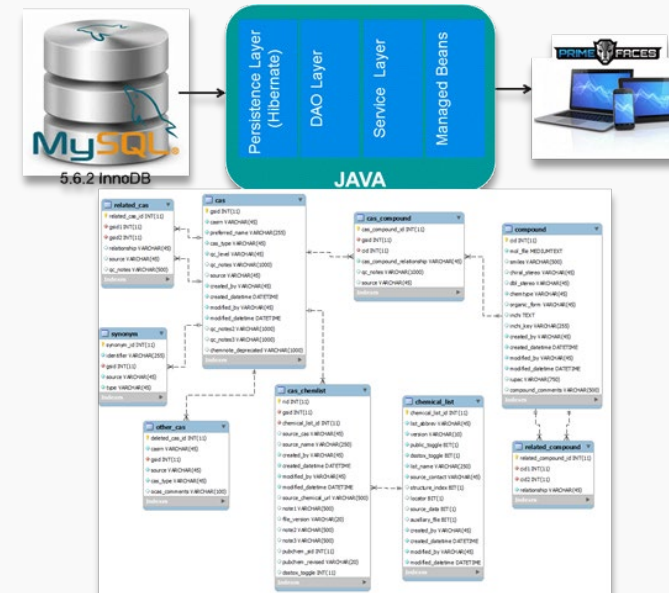


- 2005



2012

- Convert DSSTox tables to MySQL
- Develop curation interface & cheminformatics workflow
- Expand chemical content
- Web-services & Dashboard access



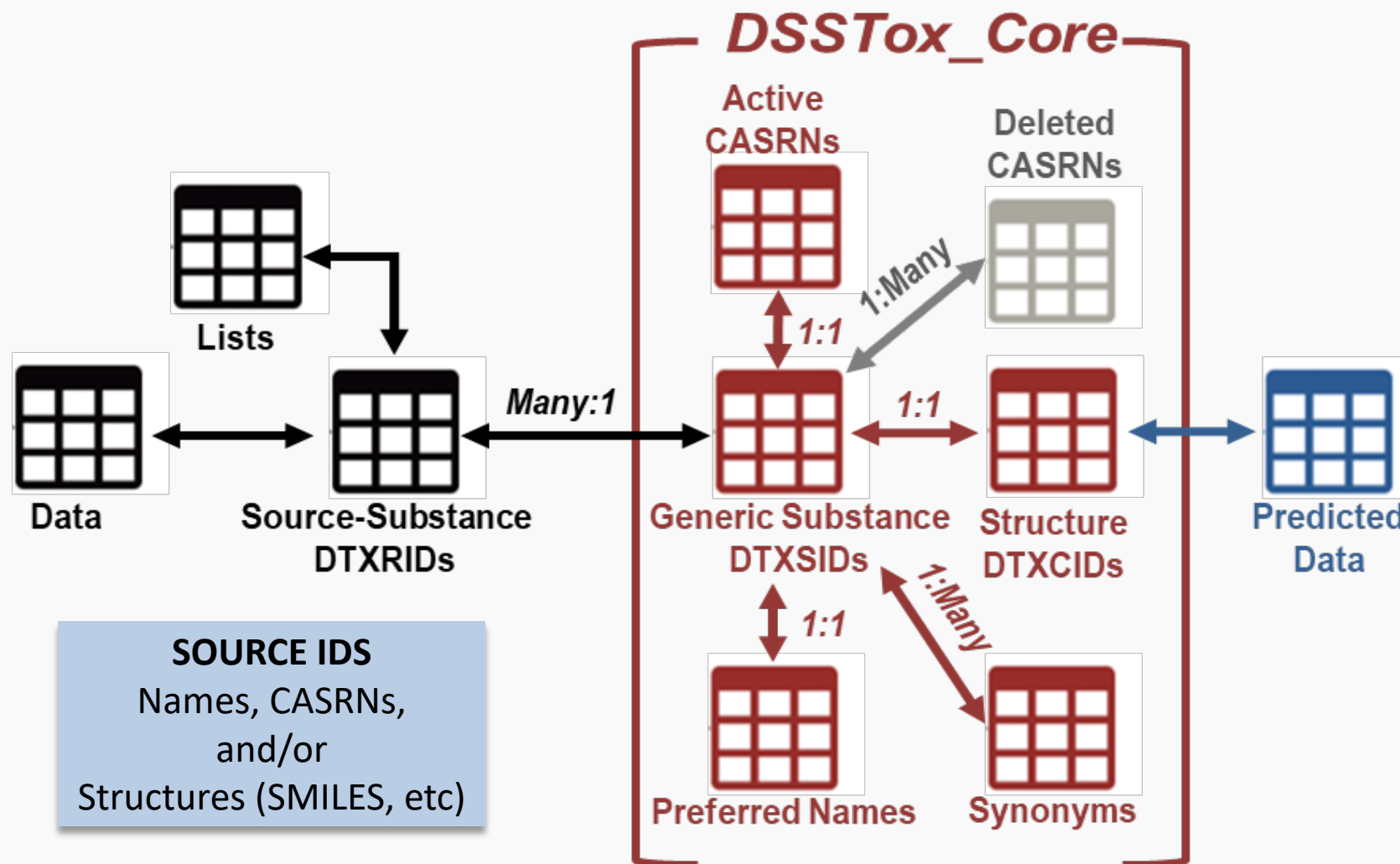
Grulke, et al., *Comput Toxicol*, 12:100096, 2019.

2014



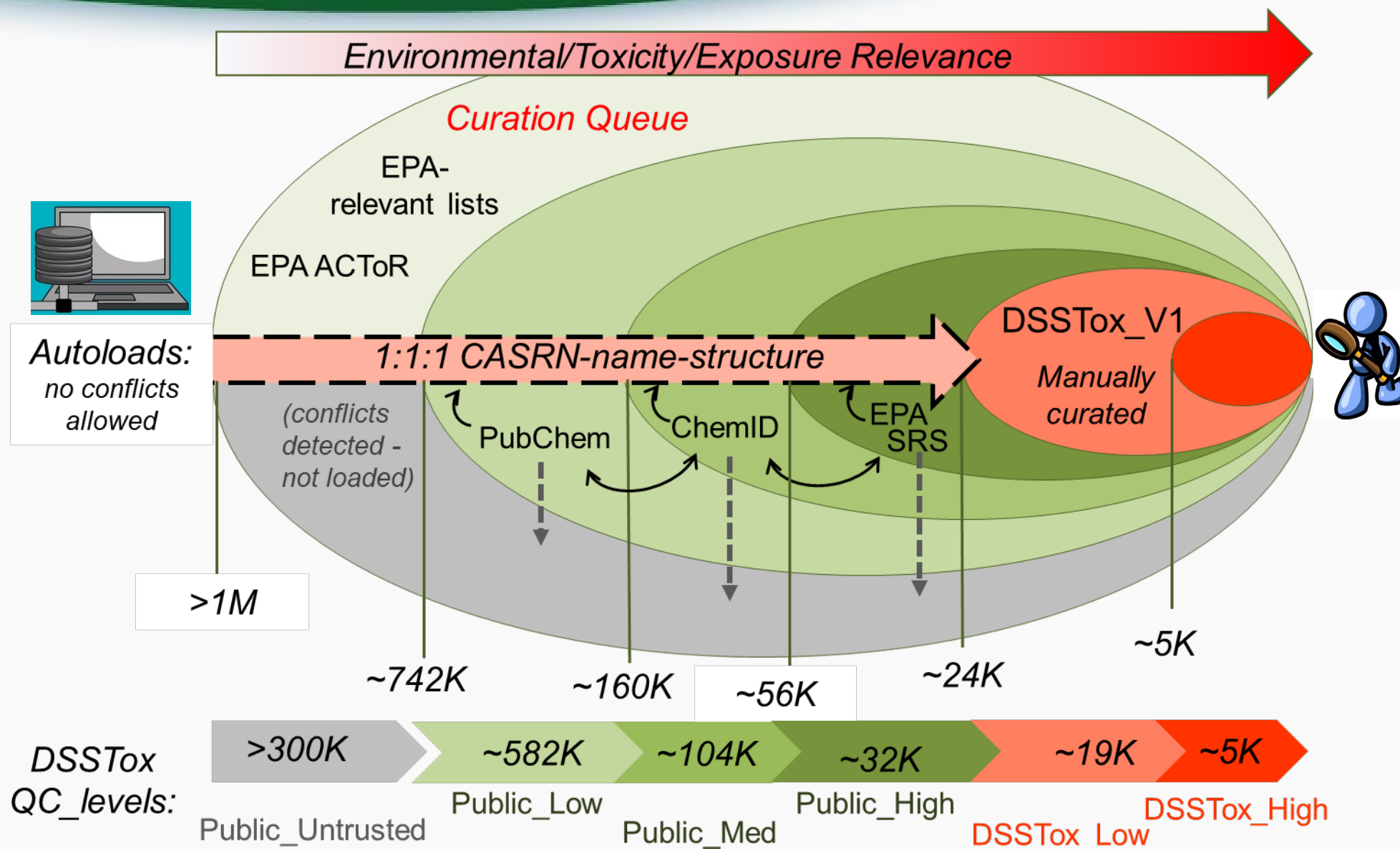
present

DSSTox Data Model



How did we go from 25K to ~1M
chemicals while addressing EPA's
needs & quality concerns?

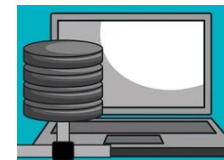
Building DSSTox_v2 (circa 2015)



Data source auto-load order:

1) DSSTox_v1 (~22K)

- ✓ 1:1 CAS-structure mappings



2) EPA SRS (~77K)

- ✓ systematic name → structure conversion
- ✓ internal CAS-structure conflicts (12.5%)
- ✓ ChemID conflicts (24% of 30K overlaps)
- ✓ DSSTox conflicts (8% of 6200 overlaps) → queue for curation



3) ChemID (~400K)

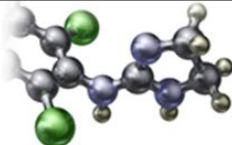
- ✓ internal CAS-structure conflicts (4.5%)
- ✓ PubChem conflicts (45% of 225K overlaps) ... OUCH!!
- ✓ DSSTox conflicts (11% of 2300 overlaps) → queue for curation



4) And so on ...

Substance-structure ambiguity

DSSTox Curation
1:1:1 CAS:Name:Structure

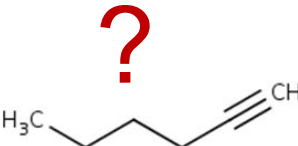


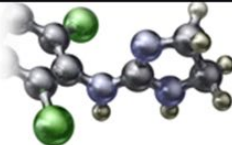
ChemIDplus
A TOXNET DATABASE

Start New Query Modify Query Search History Show Query
Switch to Summary View

Substance Name: Hexyne
RN: 26856-30-4
InChIKey: CGHIBGNXEGJPQZ-UHFFFAOYSA-N

Molecular Formula
C₆H₁₀
Molecular Weight
82.145



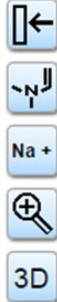
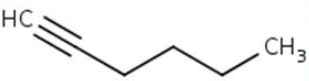


ChemIDplus
A TOXNET DATABASE

Start New Query Modify Query Search History Show Query
Switch to Summary View
Go to summary view

Substance Name: 1-Hexyne
RN: 693-02-7
InChIKey: CGHIBGNXEGJPQZ-UHFFFAOYSA-N

Molecular Formula
C₆H₁₀
Molecular Weight
82.145

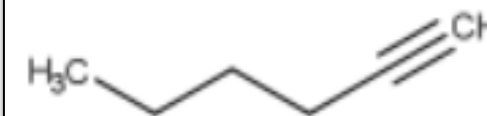


2 related chemical structures
with this substance

No Structure

Hexyne

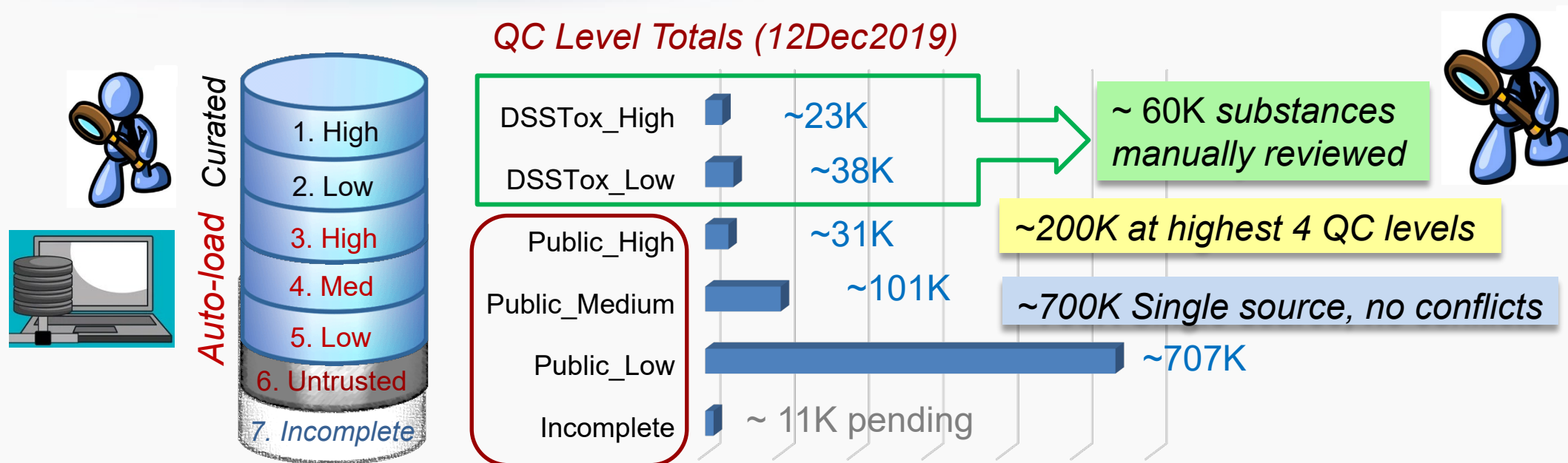
DTXSID : DTXSID601014672
CASRN : 26856-30-4



1-Hexyne

DTXSID : DTXSID30870753
CASRN : 693-02-7

DSSTox_v2 QC Levels

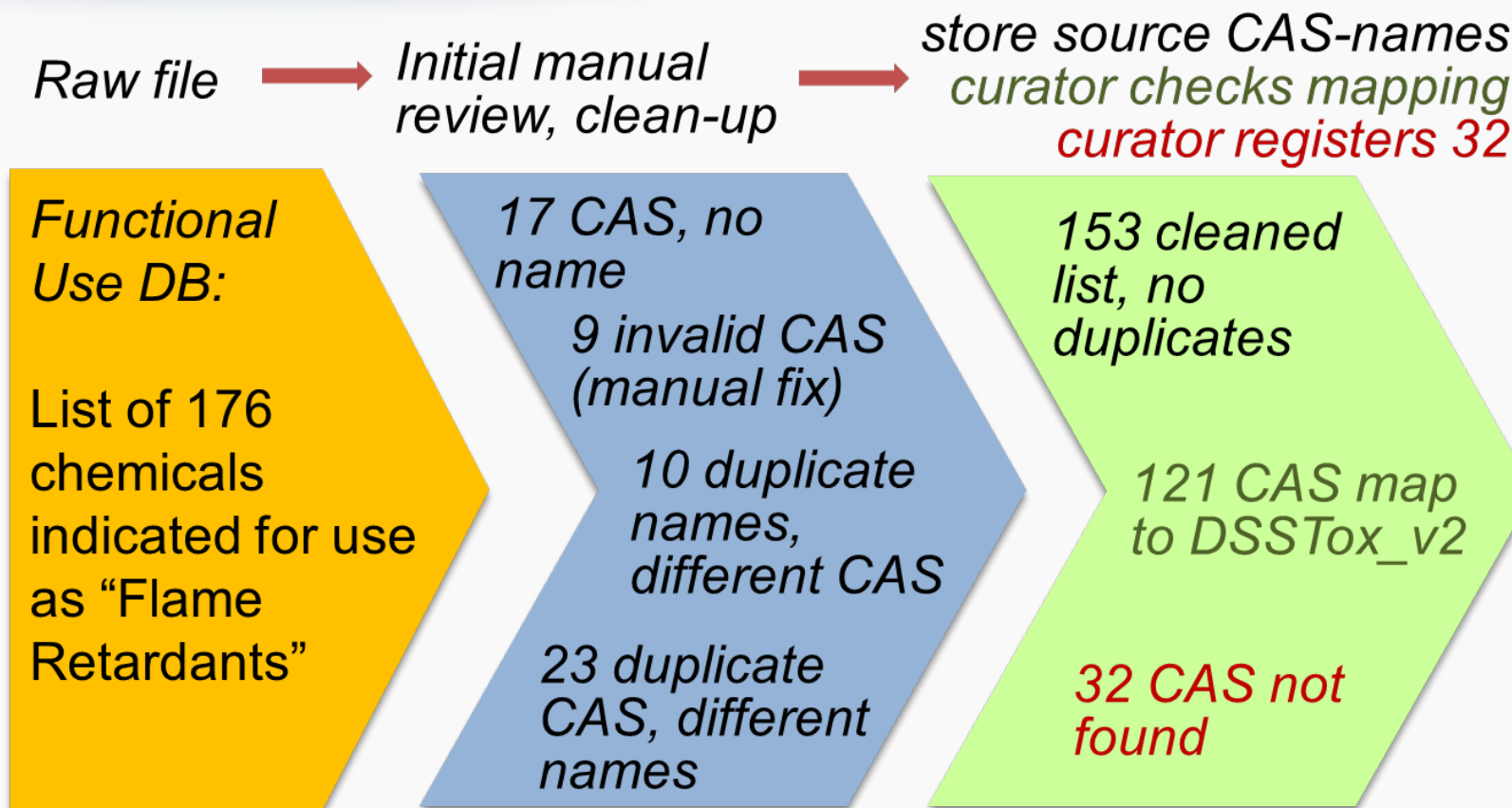


QC Levels	
DSSTox_High:	Hand curated - highest confidence from definitive source
DSSTox_Low:	Hand curated and confirmed using multiple public sources
Public_High:	Extracted from EPA SRS and confirmed to have no conflicts in ChemID and PubChem
Public_Medium:	Extracted from ChemID and confirmed to have no conflicts in PubChem
Public_Low:	Extracted from ACToR or PubChem (single source, no conflicts)
Public_Untrusted:	Postulated, but found to have conflicts in public sources

So, how and when do we manually
curate DSSTox content?



Curate list of “Flame Retardants”



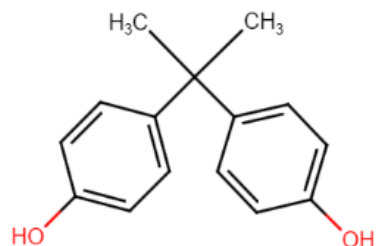
15% CAS error or no name, 19% conflicts, 18% CAS not found
80% of cleaned list map to existing DSSTox content
32 curated as new chemical registrations

DSSTox Chemical Registration/Curation Tool

ACToR-DSSTox Chemical Registration

[View/Edit a Single Record](#) [Structure Search](#) [Browse/Curate Records](#) [Export DSSTox](#) [Chemotypes](#) [Manage Chemical Lists](#) [Manage Property Data](#) [Add Deleted Casrns](#)

Preferred Name matched null
You are viewing the record associated with DTXSID7020182
CASRN: 80-05-7

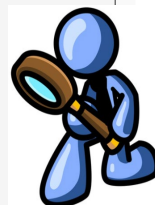

Cc1ccc(cc1)C(C)(C)c2ccc(O)cc2

Calculate from Structure

Substance_ID: DTXSID7020182
CAS:
Name:
Substance Type:
QC Level:
Data Source:
QC Notes:

Compound_ID: DTXCID30182
Chemical Shown:
Private Notes:
Source of CAS-Compound:
Double Stereo:
Chiral Stereo:
Chemical Form:
Organic Form:

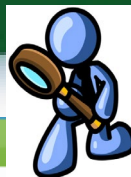
[Synonyms \(128\)](#)
[Other Cas \(6\)](#)
[Successor Substances \(17\)](#)
[Predessoror Substances \(23\)](#)
[Associated Lists \(2226\)](#)
[Save Record](#)



- ▶ Synonyms (128)
- ▶ Other Cas (6)
- ▶ Successor Substances (17)
- ▶ Predessoror Substances (23)
- ▶ Associated Lists (2226)

[Save Record](#)

DSSTox Chemical Registration/Curation Tool



Synonyms (128)

Synonyms (128)

Associated Lists (2226)

Associated Lists (2226)

Other Cas (6)

Other CAS (6)

Successor Substances (17)

Successor Substances (17)

Predecessor Substances (23)

Predecessor Substances (23)

Relationship	DTXSID	CAS-RN	Source
Predecessor: Component	DTXSID70872791	26265-08-7	STN(DSS)
Predecessor: Component	DTXSID201015347	26402-79-9	STN(DSS)
Predecessor: Component	DTXSID40105886	106906-26-7	STN(DSS)
Predecessor: Component	DTXSID4097128	68152-46-5	STN(DSS)
Predecessor: Component	DTXSID6027840	25037-45-0	STN(DSS)
Predecessor: Component	DTXSID601037135	202483-49-6	STN(DSS)
Predecessor: Component	DTXSID001037149	235420-85-6	Public
Predecessor: Component	DTXSID701037164	235420-83-4	Public
Polymer	DTXSID3049627	25085-75-0	STN(DSS)
Polymer	DTXSID0050479	25068-38-6	STN(DSS)
Polymer	DTXSID9050480	28906-96-9	STN(DSS)
Polymer	DTXSID601037164	252300-00-4	Public

DTXRID	Source Identifier	Inventory
DTXRID7020182	Bisphenol A	CPDBAS
DTXRID4022363	bisphenol A	NCTREER
DTXRID3023922	Bisphenol A.	IRISTR
DTXRID6024624	Bisphenol A	NTPGTZ
DTXRID2026492	Phenol, 4,4'-(1-methylethylidene)bis-	HPVCSI
DTXRID7030852	Bisphenol A	NTPBSI
DTXRID0033257	Bisphenol A	NTPHTS
DTXRID5033258	Bisphenol A	NTPHTS
DTXRID3034119	Bisphenol A	ICCVAM
DTXRID6036675	Bisphenol A	NCTRAR
DTXRID4037748	4,4'-Propane-2,2-diylidiphenol (bisphenol A)	KIERBL
DTXRID903978693	80-05-7	IRIS
DTXRID004834374		WIKIDATA
DTXRID204891415	80-05-7	COASummaries020920
DTXRID104924744	0000080-05-7	CPDAT
DTXRID804978661	80-05-7	toxval2
DTXRID004994736	80-05-7	CHEMINV
DTXRID805065186	80057	ECOTOX
DTXRID602677009	80-05-7	NCCTCPP_BCF
DTXRID505135671	80-05-7	EDSPuPDF
DTXRID305455617	80-05-7	DNTREF
DTXRID205459230	80-05-7	NEMILIST
DTXRID005493723	80-05-7	ARInvitro2016
DTXRID705494253	80-05-7	CTORDER1
DTXRID505494388	80-05-7	ARAntagonistRL2016d

DSSTox List Registration/Curation Tool

Create a new chemical list **Create List**

Select a chemical list **EPASPECIATE**

Edit **Add/Edit Chemicals**

Edit List **Add/Edit Chemicals**

Search for a list



ACToR-DSSTox Chemical Registration

Export DSSTox Chemotypes Manage Chemical Lists Manage Property Data Add Deleted Casrns Welcome, Ann Logout

Substance Mapping

(1 of 1) 1 25

	Source Casrn	Source Name	Hit Substance_ID	Hit Casrn	Hit Name	
•	68917-57-7	Terpene	DTXSID0029501	68917-57-7	Terpenes and Terpenoids, mixed sour and sweet Orange-oil	Other Hits
•	27323-28-0	1-Methylindole	DTXSID50950041	27323-28-0	2-Methyl-1H-indole--5-methyl-1H-indole (1/1)	Other Hits
•	7782-41-4	Fluoride ion	DTXSID3024106	7782-41-4	Fluorine	Other Hits
•	156-59-2	1,2-Dichloroethene	DTXSID2024030	156-59-2	(Z)-1,2-Dichloroethene	Other Hits
•	27813-02-1	2-Hydroxypropyl methacrylate	DTXSID5027936	27813-02-1	Hydroxypropyl methacrylate	Other Hits

Hits

	ssCAS-RN	ssName	Hit Desc	Hit Substance_ID	Hit Casrn	Hit Name
•	156-59-2	1,2-Dichloroethene	CAS-RN matched CAS	DTXSID2024030	156-59-2	(Z)-1,2-Dichloroethylene
•	156-59-2	1,2-Dichloroethene	Valid Synonym matched NAME1	DTXSID8024991	540-59-0	1,2-Dichloroethylene

Map hit **Cancel**

Identifier conflicts binned to facilitate curation

Editing Listname: EPASPECIATE
Duplicates: ☒

External Check Results	
Description	Records
CAS Unique Synonym matched other record: NAME1	4
Mapped Identifier matched NAME1	8
Preferred Name matched NAME1 CAS-RN matched other record: CAS	10
CAS-RN matched CAS Valid Synonym matched other record: NAME1	6
CAS-RN matched CAS Valid Synonym matched other record: NAME1	

Source CASRN&Names auto-mapped by:

- Preferred name
- CASRN (other CASRN)
- Synonym (valid, exact, unique, ambiguous, Source IDs)
- Name-to-Structure

2814 Total records

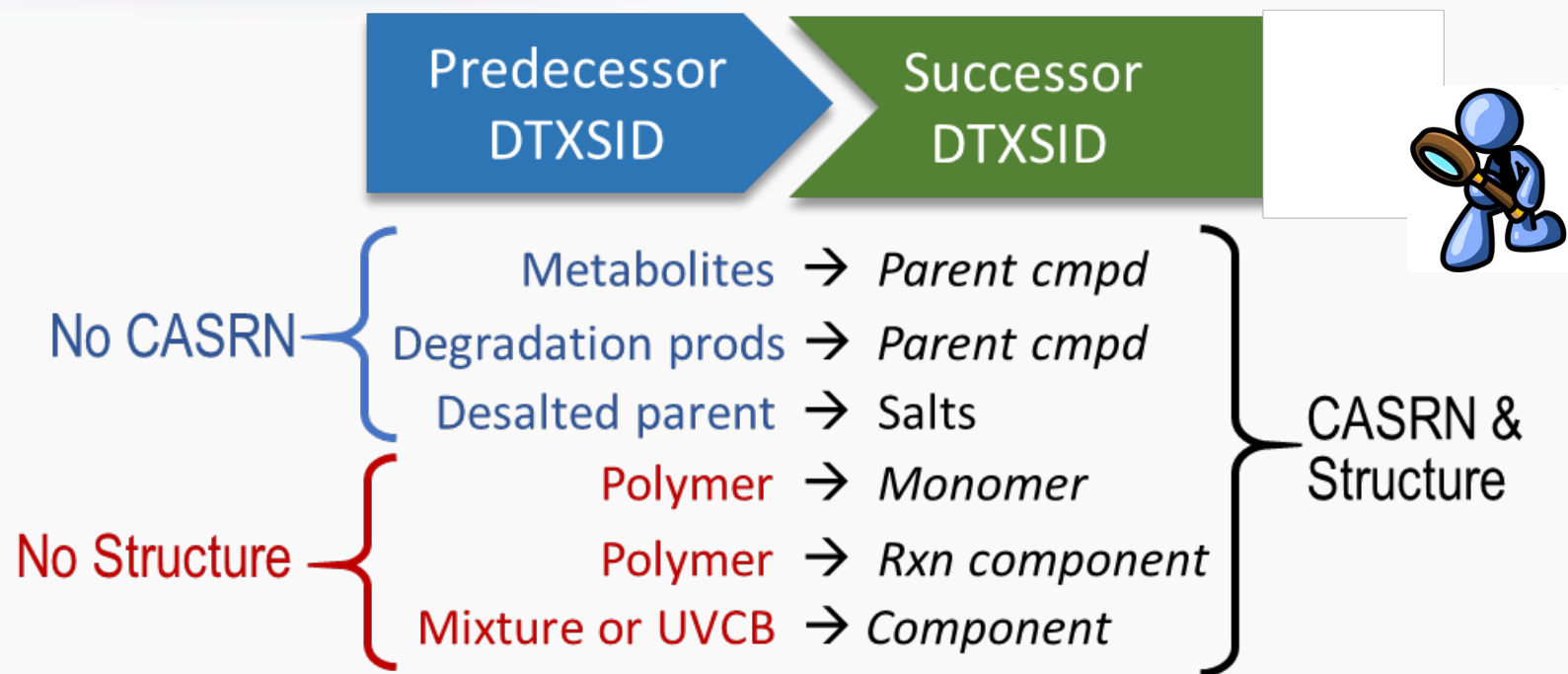
- 67 Conflict bins
- 757 No Hits

United States
Environmental Protection
Agency



M = Manual Curation Steps

Substance Relationship Mapping



- Manually added relationship tags
- Establish linkages of No CASRN → CASRN records
- Establish linkages of No Structure (e.g., UVCBs) → Structure records

No representative structures violating 1:1:1 rule

Chemical Families, e.g. PCBs

▼ Successor Substances (209)

	CAS-RN	Relationship	Source
●	32774-16-6	is a Representative Isomer of this	STN(DSSTox) ▼
●	2051-60-7	is a Representative Isomer of this	Public ▼
●	2051-61-8	is a Representative Isomer of this	Public ▼
●	2051-62-9	is a Representative Isomer of this	Public ▼
●	13029-08-8	is a Representative Isomer of this	▼
●	16605-91-7	is a Representative Isomer of this	▼
●	25569-80-6	is a Representative Isomer of this	▼
●	33284-50-3	is a Representative Isomer of this	Public ▼
●	34883-43-7	is a Representative Isomer of this	Public ▼
●	34883-39-1	is a Representative Isomer of this	Public ▼
●	33146-45-1	is a Representative Isomer of this	Public ▼
●	2050-67-1	is a Representative Isomer of this	Public ▼
●	2974-92-7	is a Representative Isomer of this	Public ▼
●	2974-90-5	is a Representative Isomer of this	Public ▼
●	34883-41-5	is a Representative Isomer of this	Public ▼
●	2050-68-2	is a Representative Isomer of this	Public ▼



<https://ccte-ced.epa.gov/dashboard/>

Related Chemicals
 Found 209 chemicals

2,4-Dichlorobiphenyl
 34883-43-7

2,2',4,4'-Tetrachlorobiphenyl
 2437-79-8

3,3',4,4'-Tetrachlorobiphenyl
 32568-13-3

2,3,4,4',5-Pentachlorobiphenyl
 31508-00-6

3,3',4,4',5-Pentachlorobiphenyl
 67465-28-8

2,2',4,4',5,5'-Hexachlorobiphenyl
 35065-27-1

2,2',3,5',6-Pentachlorobiphenyl
 38379-99-6

2,2',3,5-tetrachlorobiphenyl
 41464-39-5

2,2,4,5,5'-Pentachlorobiphenyl
 37680-73-2

Markush Representations, e.g PCBs

ACToR-DSSTox Chemical Registration

View/Edit a Single Record | Structure Search | Browse/Curate Records | Export DSSTox | Chemotypes | Manage Chemical Lists | Manage Property Data | Add Deleted Casrns

Valid Synonym matched null
You are viewing the record associated with DTXSID5024267
CASRN: 1336-36-3

PCBs

License has expired

R1 >0, Rest H

R1 =

Calculate from Structure

Substance_ID: DTXSID5024267
CAS: 1336-36-3
Name: Polychlorinated biphenyls
Substance Type: Chemical Category
QC Level: DSSTox_High
Data Source: STN(DSSTox)
biphenyl with multiple (unknown number) chlorines attached at unknown locations

Compound_ID: DTXCID301285169
Chemical Shown:

Private Notes:

Source of CAS-Compound: STN(DSSTox)
Double Stereo: None
Chiral Stereo: None
Chemical Form: Organic

Markush Query
No Structure
Tested Chemical
Markush Enumerable
Markush Query

ChemAxon
Markush Tools

CompTox Chemicals Dashboard

Welcome to the new EPA CompTox Chemicals Dashboard

The new Dashboard is a complete rebuild and is replacing the CompTox Chemicals Dashboard released on July 12th 2020.

Polychlorinated biphenyls
1336-36-3 | DTXSID5024267
Searched by Expert Validated Synonym.

Related Substances

Showing 211 of 211 chemicals

Searched Chemical:

Polychlorinated biphenyls
DTXSID: DTXSID5024267
CASRN: 1336-36-3
TOXCAS:

Predecessor: Component
2 related chemical structures with this substance
Clophen A 40
DTXSID: DTXSID6073504
CASRN: 52306-92-8
TOXCAS:

Markush Child

2-Chlorobiphenyl
DTXSID: DTXSID6040298
CASRN: 2051-60-7

Markush Child

3,4-Dichlorobiphenyl
DTXSID: DTXSID6073310
CASRN: 2074-92-7

Markush Child

2,2',5-Trichlorobiphenyl
DTXSID: DTXSID6073491
CASRN: 97690-65-2
TOXCAS:

Markush Child

2,2',3,4,5'-Pentachlorobiphenyl
DTXSID: DTXSID6073497
CASRN: 99390-02-8
TOXCAS:

Markush Child

2,2',3,3',6,6'-Hexachlorobiphenyl
DTXSID: DTXSID6073499
CASRN: 99411-22-2
TOXCAS:

Markush Child

2,3',4,5,5'-Pentachlorobiphenyl
DTXSID: DTXSID6073609
CASRN: 68194-12-7
TOXCAS:

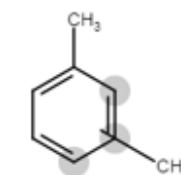
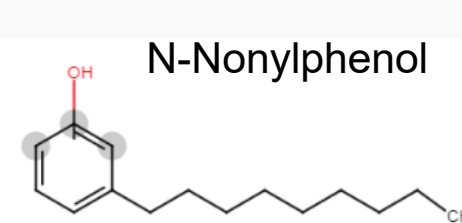
209 "Markush Children"

EPA
United States
Environmental Protection
Agency

Alcohols, C6-12, ethoxylated

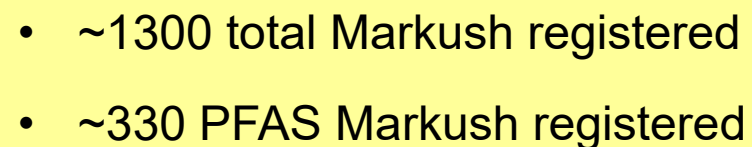
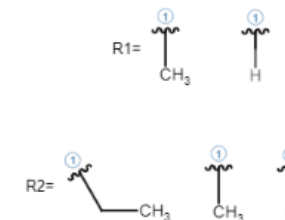
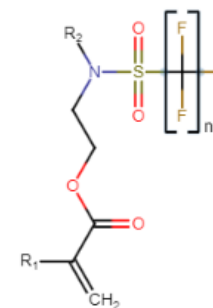


Xylenes



(N-Alkylperfluoroalkanesulfonamido)ethyl acrylate

- 



Impact of Expert Chemical Curation, e.g., PFAS



EPA's PFAS Action Plan — Released February 14, 2019

2001: first scientific study of PFOA in the global environment
2000-2002: voluntary phase-out of PFOA and PFOS

Science Inventor

You are here: EPA Home » Science Inventor

DEVELOPMENT

Citation:

LAU, C. S. DEVELOPMENTAL TOXICITY OF PERFLUOROOCTANOIC ACID (PFOA) IN THE 8th Annual National Forum on Chemical Safety, September 10-12, 2007, EPA, Washington, DC, 2007 Feb; 9(2):1-10.

Developmental toxicity

Perfluorooctanoic acid (PFOA) is a synthetic chemical used in a wide range of commercial products. PFOA is a persistent organic pollutant (POP) and is known to be toxic to aquatic life. PFOA is also a potential human health concern. PFOA is a synthetic chemical used in a wide range of commercial products. PFOA is a persistent organic pollutant (POP) and is known to be toxic to aquatic life. PFOA is also a potential human health concern.

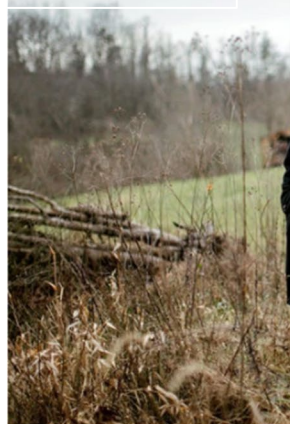
Author information

Abstract
Perfluorooctanoic acid (PFOA) is a synthetic chemical used in a wide range of commercial products. PFOA is a persistent organic pollutant (POP) and is known to be toxic to aquatic life. PFOA is also a potential human health concern. PFOA is a synthetic chemical used in a wide range of commercial products. PFOA is a persistent organic pollutant (POP) and is known to be toxic to aquatic life. PFOA is also a potential human health concern.

The Story Behind the E.P.A.'s Contaminated Water Revelation

<https://www.nytimes.com/2016/05/27/us/politics/the-story-behind-the-e-p-a-s-contaminated-water-revelation.html>

May 27, 2016



Rob Bilott near Parkersburg, W.Va., in December 2015.

By Nathaniel Rich

May 27, 2016

Last week 5.2 million Americans learned that the water in their homes was contaminated with man-made chemicals. The U.S. Environmental Protection Agency (EPA) announced that it had found perfluorooctanoic acid (PFOA) and perfluorooctane sulfonic acid (PFOS) in the drinking water of several communities in North Carolina. The chemicals are man-made and have been found in a wide range of commercial products. The EPA is working to identify the sources of the contamination and to take steps to reduce the levels of the chemicals in the water.

ENVIRONMENTAL Science & Technology LETTERS

Nov, 2016

Legacy and Emerging Perfluoroalkyl Substances in Drinking Water Contaminants in the Cape Fear River, North Carolina

Mei Sun,^{1,2,3,4} Elisa Arevalo,² Mark Strynar,⁵ Andrew Lindstrom,⁶ Adam Pickett,¹ Chris Smith,⁷ and Detlef R. U. Knappe²

¹Department of Civil and Environmental Engineering, University of North Carolina at Chapel Hill, United States

²Department of Civil, Construction, and Environmental Engineering, North Carolina State University, 27695, United States

³National Exposure Research Laboratory, U.S. Environmental Protection Agency Research Triangle Park, United States

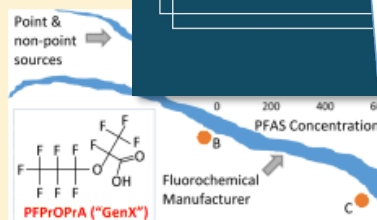
⁴Cape Fear Public Utility Authority, Wilmington, North Carolina 28403, United States

⁵Town of Pittsboro, Pittsboro, North Carolina 27312, United States

⁶Fayetteville Public Works Commission, Fayetteville, North Carolina 28301, United States

Supporting Information

ABSTRACT: Long-chain per- and polyfluoroalkyl substances (PFASs) are being replaced by short-chain PFASs and fluorinated alternatives. For ten legacy PFASs and seven recently discovered perfluoroalkyl ether carboxylic acids (PFEACs), we report (1) their occurrence in the Cape Fear River (CFR) watershed, (2) their fate in water treatment processes, and (3) their adsorbability on powdered activated carbon (PAC). In the headwater region of the CFR basin, PFEACs were not detected in raw water of a drinking water treatment plant (DWTP), but concentrations of legacy PFASs were high. The U.S. Environmental Protection Agency's (EPA) lifetime health advisory level (70 ng/L) for perfluorooctanoic acid (PFOA) was exceeded in the raw water. The EPA is working to identify the sources of the contamination and to take steps to reduce the levels of the chemicals in the water.



Feb 14, 2019

EPA Takes Aim at PFAS

PFAS, or poly- and perfluoroalkyl substances, are a wide range of persistent organic pollutants. They are used in a wide range of commercial products. PFAS are persistent organic pollutants (POPs) and are known to be toxic to aquatic life. PFAS are also a potential human health concern. PFAS are synthetic chemicals used in a wide range of commercial products. PFAS are persistent organic pollutants (POPs) and are known to be toxic to aquatic life. PFAS are also a potential human health concern.

EPA United States Environmental Protection Agency



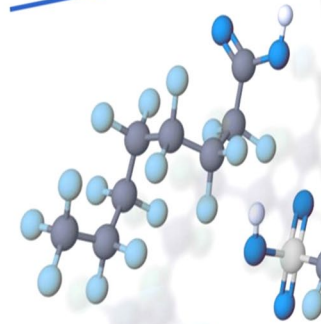
EPA's Per- and Polyfluoroalkyl Substances (PFAS) Action Plan

4 years

It can take up to 4 years for the level of PFAS in the body to go down

4,000

PFAS is a group of more than 4,000 very stable synthetic chemicals



EPA United States Environmental Protection Agency

Per- and Polyfluoroalkyl Substances (PFAS)

EPA's PFAS Action Plan

February 14, 2019 - EPA's PFAS Action Plan responds to extensive public interest and input the agency has received over the past year and represents the first time EPA has built a multi-media, multi-program, national communication and research plan to address an emerging environmental challenge like PFAS. [Read more.](#)

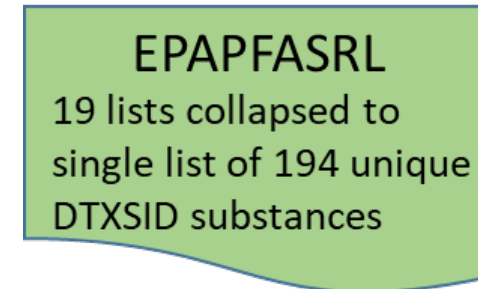
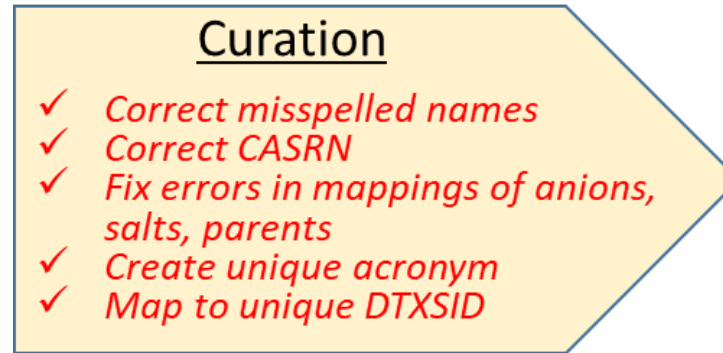
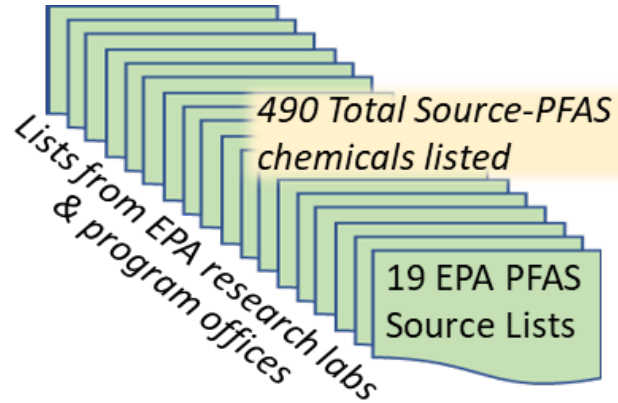
"We are moving forward with several important actions, including the maximum contaminant level process, that will help affected communities better monitor, detect, and address PFAS."

— EPA Administrator, Andrew Wheeler

<https://www.epa.gov/pfas/epas-pfas-action-plan>

PFAS: EPA Cross-Agency Research List

https://comptox.epa.gov/dashboard/chemical_lists/epapfasrl



DTXSID	Substance_Name	Substance_CASRN	Source_Name (incorrect or ambiguous)	Source_CASRN (incorrect or invalid)	Source_Acronym (incorrect or ambiguous)	Unique_Acronym
DTXSID20874028	2H,2H,3H,3H-Perfluorooctanoic acid	914637-49-3	5:3 Polyfluorinated acid	914637-49-3	5:3 acid	5:3 PFOA
DTXSID7027831	N-Methyl-N-(2-hydroxyethyl)perfluorooctanesulfonamide	24448-09-7	N-Methyl perfluorooctanesulfonam		NMeFOSE, MeFOSE	NMeFOSE
DTXSID10892352	Perfluoro-2-[(perfluoro-3-(perfluoroethoxy)-2-p	749836-20-2	Ethanesulfonic acid, 2-[1-(tetrafluoroethoxy)methyl]-1,2,2,2-tetrafluoroethoxy]-1,1,2,2-tetrafluoro		PFESA Byproduct 2	PFESA Byproduct 2
DTXSID70892479	Perfluoropentanesulfonate	270505-36-9	Pe	2706-91-4	PFPeS	PFPeS_ion
DTXSID1354	Ammonium perfluoropentanesulfonate	68259-09-6	Am	68259-09-6		APFPeS
DTXSID1081350	4,8-Dioxa-3H-perfluorononanoic acid	919005-14-4	2,2,3-trifluoro-3-[(1,1,2,2,3,3-hexafluoro-3-(trifluoromethoxy)propoxy)propanoic acid	919005-14-4	ADONA	ADONA parent acid
DTXSID00874026	Ammonium 4,8-dioxa-3H-perfluorononanoate	958445-44-8	Ammonium 4,8-dioxa-3H-pe	958445-44-8	ADONA	ADONA
DTXSID3037707	Potassium perfluorobutanesulfonate	29420-49-3	Po		PFBS	PFBS-K
DTXSID5030030	Perfluorobutanesulfonic acid	375-73-5	Pe	375-73-5	PFBS	PFBS
DTXSID60873015	Perfluorobutanesulfonate	45187-15-3	Perfluorobutanesulfonate	375-73-5	PFBS	PFBS_ion
DTXSID3040148	Perfluorodecanesulfonic acid	335-77-3	Perfluorodecanesulfonic acid		PEDS	PEDS

Anions

Misspelled names

Incorrect CAS

Ambiguous acronyms

"5:3 acid"

PFAS Name-Structure Curation Challenges

- Name too short – what the heck is it??

➤ Surflon S 382, 5:3 Acid, ? (Fluowet® PL 80B)

- Long systematic names obscure perfluoro aspect

1,1,1,2,2,3,3,4,4,5,5,6,6,7,7,8,8-Heptafluorodecane → (Perfluoro-n-octyl)ethane

- Long names difficult to use or decipher to structures

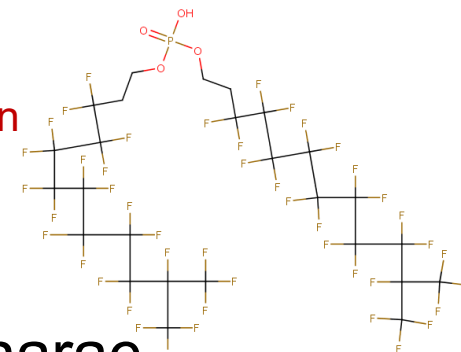
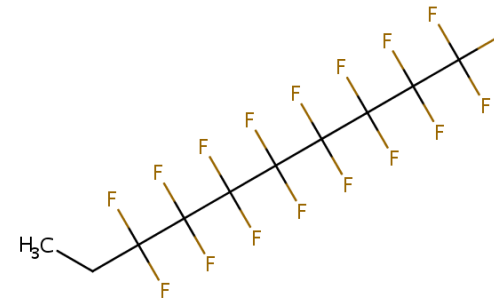
1-Dodecanol, 3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,11,12,12,12-eicosafluoro-11-(trifluoromethyl)-, hydrogen phosphate

→ Bis((perfluoro-9-methyldecyl)ethyl) phosphate

- Names so long (>256 charac) they truncated on import, incl. funky charac

2-Propenoic acid, 3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,11,11,12,12,12-heneicosafluorododecyl ester, polymer with 3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,10-heptafluorodecyl 2-propenoate, -(2-methyl-1-oxo-2-propenyl)- α ë-(2-methyl-1-oxo-2-propenyl)oxypoly(oxy-1,2-ethanediyl), 3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,11,11,12,12,13,13,14,14,15,15,16,16,16-nonacosafuorohexadecyl 2-propenoate, octadecyl 2-propenoate and 3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,11,11,12,12,13,13,14,14,14-pentacosafuorotetradecyl 2-propenoate

→ DTXSID40882569, 119973-84-1 (name too long)



Translating Expert Categories to Markush



Expert category

Fluorotelomer acrylates

Fluorotelomer alcohols

Polyfluorinated alcohols

Fluorotelomer sulfonates

N-alkyl perfluoroalkyl sulfonamidoacetic acids

N-alkyl perfluoroalkyl sulfonamidoethanols

Perfluoroalkyl aldehydes

Perfluoroalkyl amides

Perfluoroalkyl carboxylates

Perfluoroalkyl acyl fluorides

Perfluoro vinyl esters

Perfluoroalkyl ketones

Semi-fluorinated alkenes

Perfluoroalkyl vinyl ethers

Perfluoroalkyl alkyl ethers

Fluorotelomer amines

Perfluoroalkyl sulfonamides

Semi-fluorinated alkanes

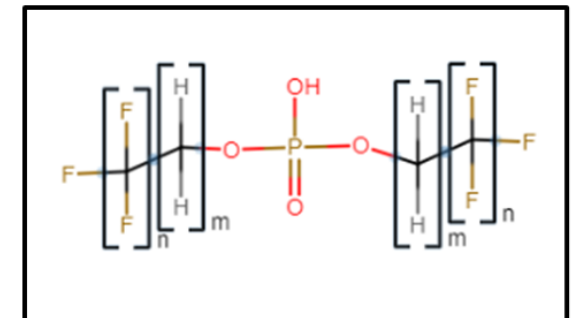
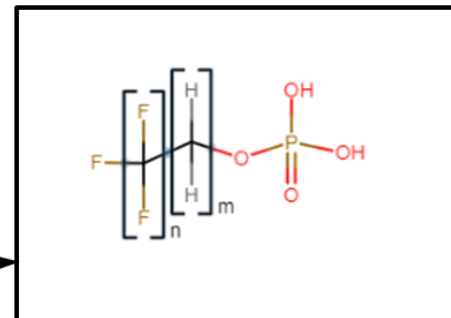
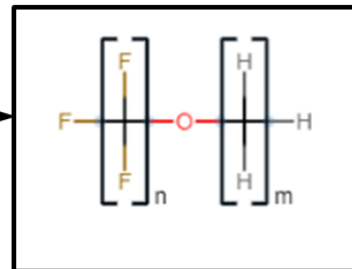
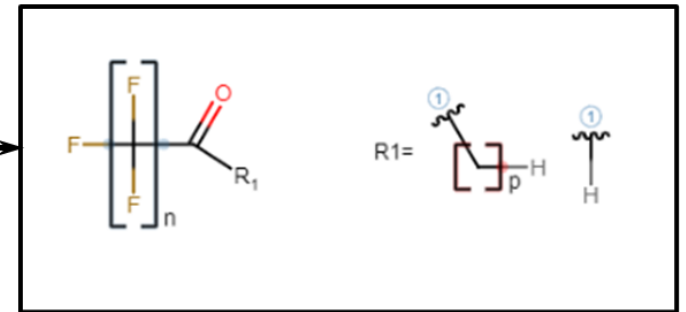
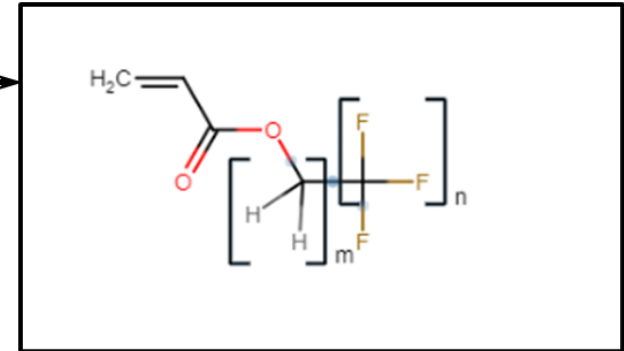
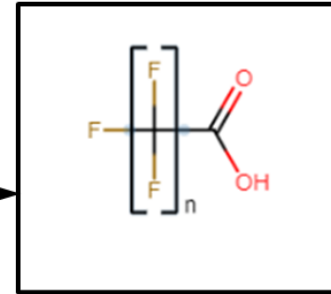
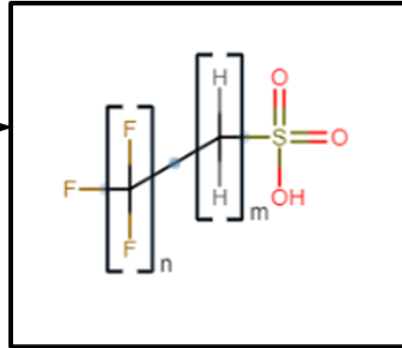
Perfluoroalkyl sulfonates

Perfluoroalkyl sulfonamido amines

Polyfluoroalkyl carboxylates

Perfluoroalkyl ethers

Fluorotelomer phosphates



PFAS Lists on the Dashboard

https://comptox.epa.gov/dashboard/chemical_lists/?search=PFAS

Home Search Lists About Tools Submit Comments Search all data

Welcome to the new EPA CompTox Chemicals Dashboard
complete rebuild and is replacing the CompTox Chemicals Dashboard released on July 12th 2020.

Lists of Chemicals
List of Assays

Chemical Lists

PFAS

39/323 public lists are PFAS-related

EXPORT COPY URL

PFASMASTER (12034) list includes:

- 10776 structures (PFASSTRUCT)
- 1258 Non-structurables (UVCBs, polymers, mixtures, etc.)
- Covers all PFAS lists (e.g., PFASOECD)

Assembly and Curation of Lists of Per- and Polyfluoroalkyl Substances (PFAS) to Support Environmental Science Research

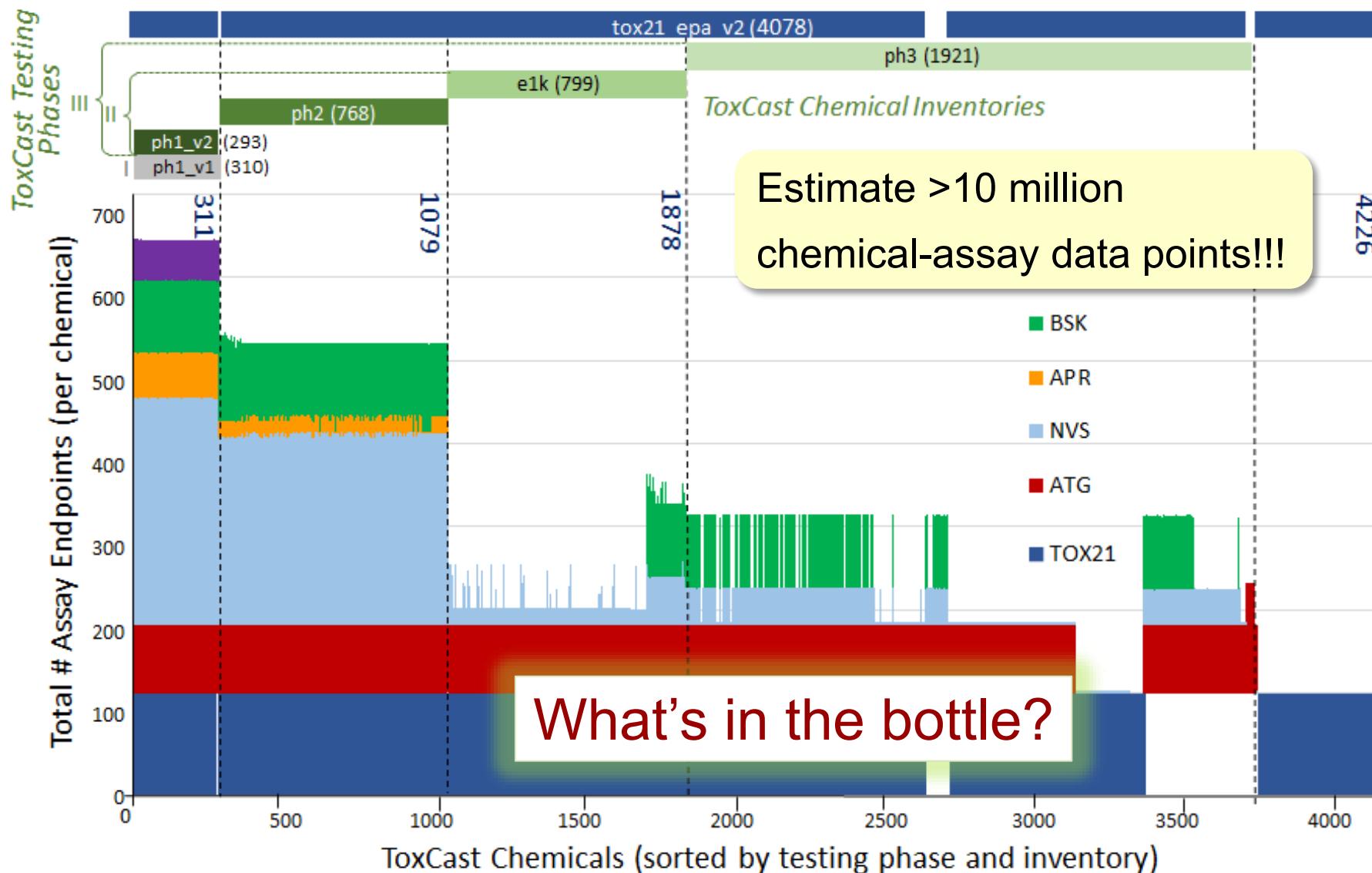
Antony J. Williams^{1*}, Linda G. T. Gaines², Christopher M. Grulke^{1†}, Charles N. Lowe¹, Gabriel F. B. Sinclair³, Vicente Samano⁴, Inthirany Thillainadarajah⁴, Bryan Meyer⁴, Grace Patlewicz² and Ann M. Richard¹

Front. Environ. Sci., 2022
<https://doi.org/10.3389/fenvs.2022.850019>

Impact of Expert Chemical Curation, cont.



ToxCast/Tox21 High-throughput Screening



ChemTrack:

Chemical Sample Management

chemtrack Inventory ▾ Partnerships ▾ Ordering ▾ Data

Single Chemical Search 🔍

Ann ▾

Bottle Inventory

Add New

Uploaded MOSAIC Files 48 **Curated Bottles 53607** Uncurated Bottles 3360

Columns ▾ 10 ▾

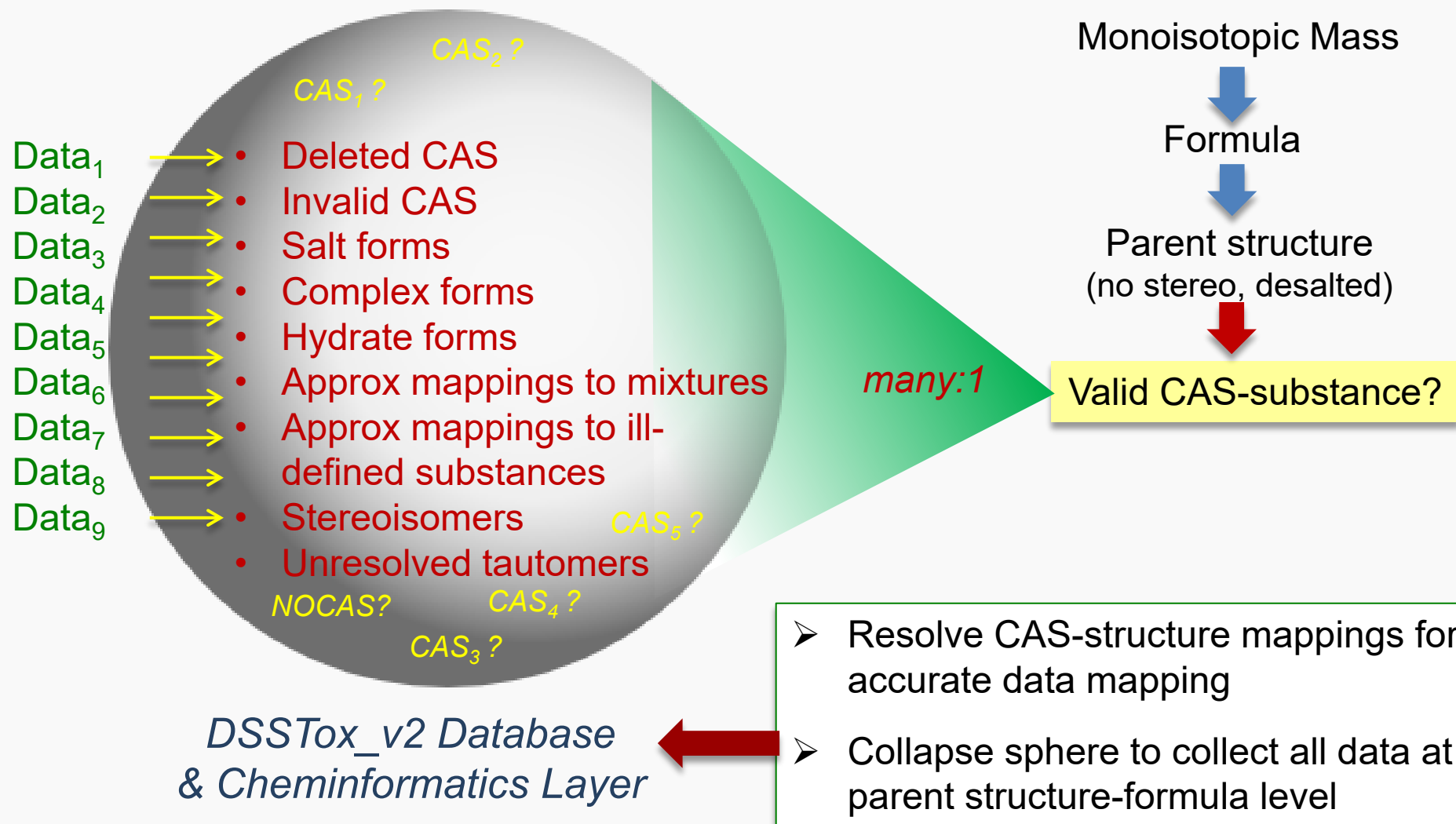
Barcode	CoA Summary Id	Barcode type	DTXSID	Compound name	Casrn	Qty available	Concentration (mM)	Vendor
Filter by Barc	Filter by CoA	Filter by Barcode ty	Filter by DTXSI	Filter by Compou	Filter	Filter	Filter by Co	Filter
0073019474	7083	EPA_Vial_MLR96_(1.4mL)	DTXSID2047648	3-PHENYLPROPYL ACETATE	122-72-5	1000ul	20	Sigma Chemical Company
0073019475	7065	EPA_Vial_MLR96_(1.4mL)	DTXSID6047604	PHENETHYL BUTYRATE	103-52-6	955ul	20	Sigma Chemical Company
0073019476	7089	EPA_Vial_MLR96_(1.4mL)	DTXSID5021621		116-53-0	1000ul	20	Sigma Chemical Company

- DSSTox curators manually review COAs & MSDSs and map each original supplier bottle to a DTXSID
- Neat and solution tracking for >53K bottle codes!

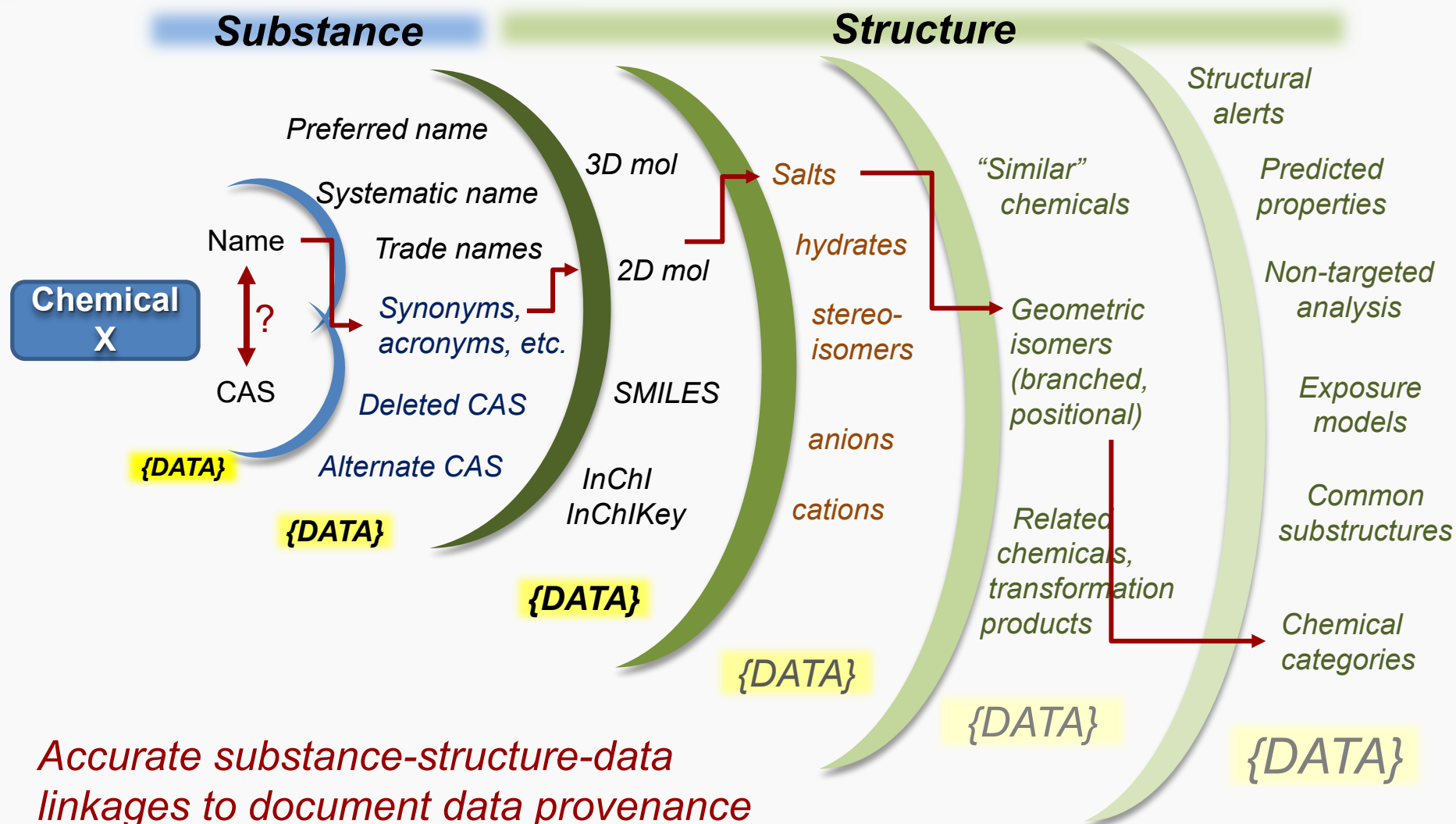


Non-targeted MS screening of environmental samples

CAS-Name-Structure “Sphere of Confusion”



Chemical-data gathering for systematic reviews



Chemical-data gathering for systematic reviews: e.g., PFOS

Perfluorooctanesulfonic acid <https://comptox.epa.gov/dashboard/dsst/substances?search=DTXSID3031864>

1763-23-1 | DTXSID3031864

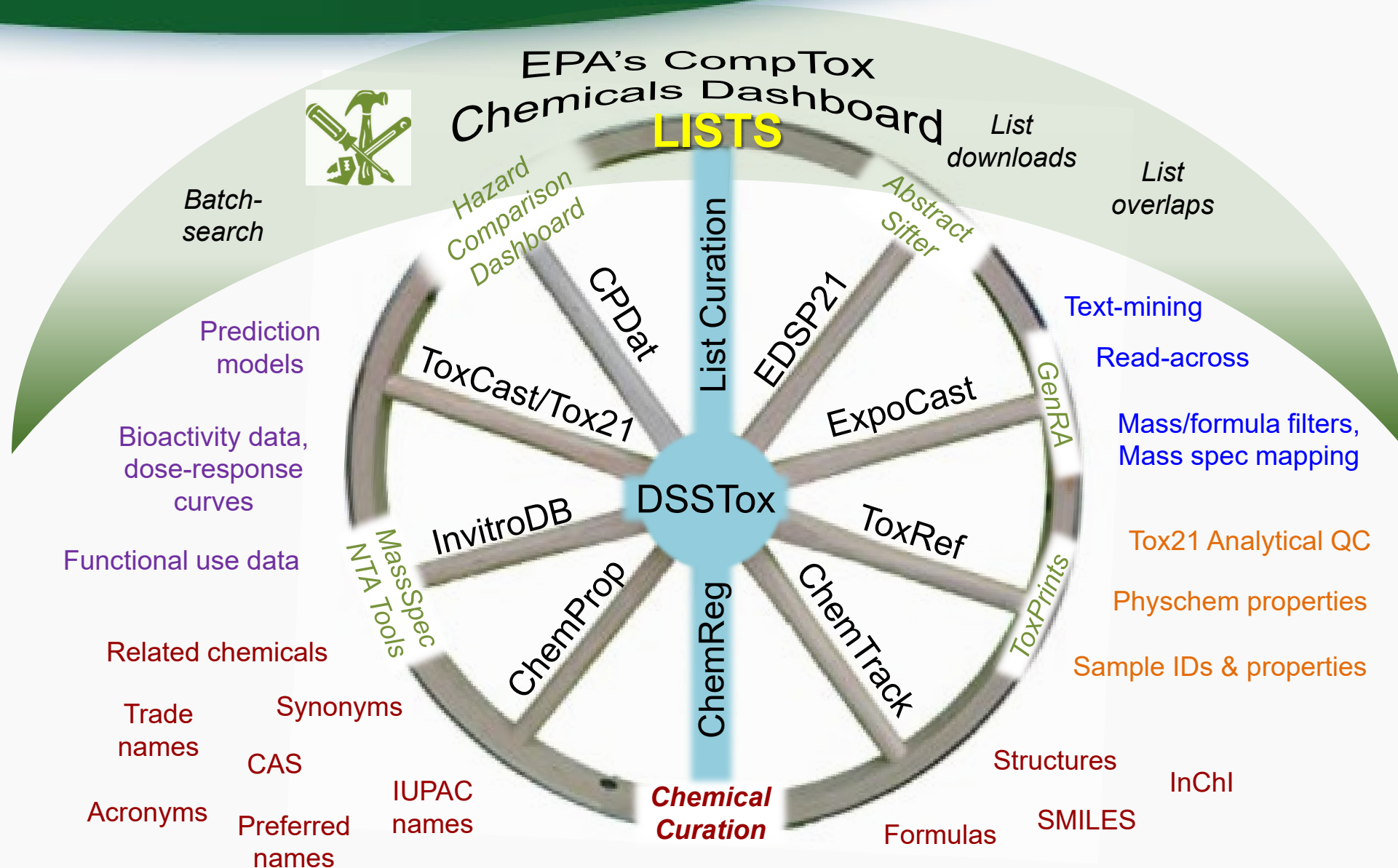
Perfluorooctanesulfonic acid

1763-23-1 | DTXSID3031864

Searched by DSSTox Substance Id.

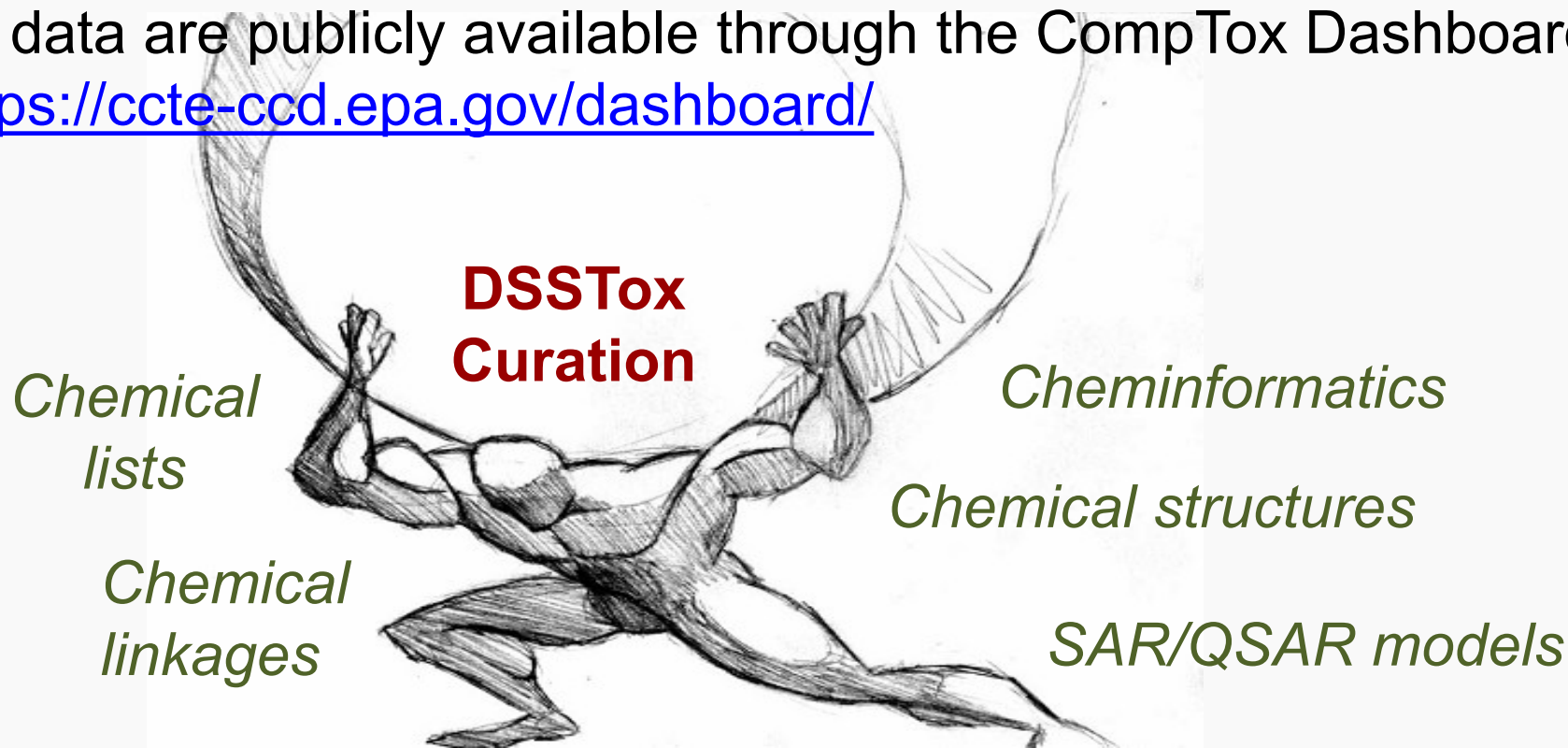
General	Toxicology	Publications	Analytical	Prediction
- EPA Substance Registry Service - PubChem - Chemspider - CPCat - DrugBank - Wikipedia - MSDS Lookup - ChEMBL - ToxPlanet - ACS Reagent Chemicals - Wolfram Alpha - ECHA Infocard - ChemAgora - Consumer Product Information Database - ChEBI - NIST Chemistry Webbook - WEBWISER - PubChem Safety Sheet - Consumer Product Information Database - PubChem: Chemical Vendors - CAMEO Chemicals - State-Specific Water Quality Standards - Wikidata - ECHA Brief Profile - NIOSH Chemical Safety Cards - PubChem 3D conformer download - EPA Hazard Comparison Dashboard - PubChem 3D Structure Display - CAS Common Chemistry - Scholia Wikidata	- ACToR - DrugPortal - CCRIS - ChemView - CTD - eChemPortal - Gene-Tox - HSDB - ACToR PDF Report - CREST - National Air Toxics Assessment - ChemView - Chemical Checker - BindingDB - CalEPA OEHA - NIOSH IDLH Values - LactMed - ECOTOX	- Toxline - PPRTVWEB - PubMed - IRIS Assessments - EPA HERO - NIOSH Skin Notation Profiles - NIOSH Pocket Guide - RSC Publications - BioCaddie DataMed - Springer Materials - Bielefeld Academic Search Engine - CORE Literature Search - Google Books (Text Search) - Google Patents (Text search) - Google Scholar (Text search) - Google Patents (Structure search) - Google Books (Structure Search) - Google Scholar (Structure search) - Federal Register	- RSC Analytical Abstracts - Tox21 Analytical Data - MONA: MassBank North America - mzCloud - NIST IR Spectrum - NIST MS Spectrum - MassBank - NIST Antoine Constants - IR Spectra on PubChem - NIST Kovats Index values - Protein DataBank - National Environmental Methods Index	2D NMR HSQC/HMBC Prediction - Carbon-13 NMR Prediction - Proton NMR Prediction - ChemRTP Predictor - LSERD

External Links



Conclusions

- DSSTox has gained a reputation for quality chemical substance-structure content supporting a wide range of EPA programs
- Strict data model constraints (1:1:1) and expert manual curation are critical components
- All data are publicly available through the CompTox Dashboard
<https://ccte-cdd.epa.gov/dashboard/>



EPA's Senior Environmental Employment (SEE) Program

DSSTOX-884



Register No Hits in
SWGDRUGv2 (228
NOHITS) - Batch 4

Opioids, Cannabin...



DSSTOX-877



Register No Hits in
SWGDRUGv2 (1342
NOHITS) - Batch 2

Opioids, Cannabin...



DSSTOX-885



Register No Hits in
SWGDRUGv2 (182
NOHITS) - Batch 5

Opioids, Cannabin...



DSSTOX-888



Curate 10 NOHITS
for the MTx700
metabolic

Metabolites and M...



Inthirany (Indira) Thillainadarajah

BA/MS Chemistry,
>15 yrs EPA laboratory & DSSTox curation
experience

Brian Meyer

PhD Medicinal & Pharmacognosy Chemistry
(Bayer Crop Sciences, retired)

Vicente Samano

PhD Organic Chemistry
(GlaxoSmithKline, retired)

Sakuntala (Saku) Sivasupramaniam

PhD Entomology & Plant Sciences (Monsanto,
retired)

Acknowledgements:

EPA's Incredible DSSTox Chemistry Team

Chris Grulke (formerly EPA- we miss him terribly!)

Current position - Assoc. Dir. Cheminformatics, Wave Life Sciences



Antony Williams, EPA's Chemistry & Cheminformatics Branch



Gabriel Sinclair, ORAU Student Services Contractor to US EPA



Indira Thillainadarajah
Brian Meyer
Saku Sivasupramaniam
Vicente Samano

} EPA's Senior Environmental
Employment (SEE) Program

This work was reviewed by EPA and approved for publication but does not necessarily reflect official Agency policy. Mention of trade names or commercial products does not constitute endorsement.